

## DOCUMENT RESUME

ED 266 165

TM 860 109

**AUTHOR** Stiggins, Richard J.; And Others  
**TITLE** A Feeling for the Student: An Analysis of the Art of Classroom Assessment.  
**INSTITUTION** Northwest Regional Educational Lab., Portland, OR. Center for Performance Assessment.  
**SPONS AGENCY** National Inst. of Education (ED), Washington, DC.  
**PUB DATE** 30 Nov 85  
**CONTRACT** 400-85-0005  
**NOTE** 173p.  
**PUB TYPE** Reports - Research/Technical (143)

**EDRS PRICE** MF01/PC07 Plus Postage.  
**DESCRIPTORS** Academic Achievement; Case Studies; \*Classroom Observation Techniques; Classroom Techniques; \*Educational Testing; Elementary Secondary Education; Evaluation Methods; Grade 6; \*Informal Assessment; Literature Reviews; Self Evaluation (Individuals); \*Student Evaluation; Teacher Characteristics; \*Teacher Made Tests; \*Teacher Role

**IDENTIFIERS** Journal Writing

**ABSTRACT**

Conducted to define in detail the ingredients contributing to an effective classroom assessment environment, this study explored the assessment methods actually used by classroom teachers through case study observations and teacher journals. Chapter 1 provides an overview of the report. Chapter 2 presents the findings of a review of the literature on educational testing, teacher decision making, and the ethnographics of classroom assessment which formed the basis for the research design. (A seven page bibliography is appended to this chapter.) Chapter 3 presents three ethnographic studies (out of the 30 case studies conducted) of 6th grade classrooms in which researchers were participant observers for 20 days. Chapter 4 presents an analysis of assessment journals maintained by 32 elementary and secondary teachers over a 10-week period. Chapter 5 summarizes the findings. It concludes that while the assessment environment is under the direct control of the teacher, there are six sets of factors that constrain teachers in establishing that environment effectively: (1) teacher characteristics; (2) classroom characteristics; (3) school and district policy; (4) instructional materials available; (5) specific school subject characteristics; and (6) the actual assessments used. The specific ingredients of these constraining factors are listed along with the relevant underlying continuum and recommended observation strategies. (BS)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

ED266165

TM 860 109

**A FEELING FOR THE STUDENT:**  
An analysis of the art of classroom assessment

Richard J. Stiggins  
Nancy Bridgeford  
Nancy Conklin  
and  
Celeste Brody

U.S. DEPARTMENT OF EDUCATION  
NATIONAL INSTITUTE OF EDUCATION  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

☒ This document has been reproduced as  
received from the person or organization  
originating it.

☐ Minor changes have been made to improve  
reproduction quality.

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official NIE  
position or policy.

Center for Performance Assessment  
Northwest Regional Educational Laboratory  
300 S.W. Sixth Avenue  
Portland, Oregon 97204

Deliverable Due November 30, 1985  
Final Research Report

The work upon which this publication is based was performed pursuant to  
Contract 400-83-0005 of the National Institute of Education. It does  
not, however, necessarily reflect the views of that agency.



## Table of Contents

|                                                      | <u>Page</u> |
|------------------------------------------------------|-------------|
| Chapter 1      Overview of the Report                | 1           |
| Chapter 2      Insights Into Classroom Assessment    | 6           |
| The Measurement Paradigm                             | 7           |
| Insights from Research on Testing                    | 11          |
| Insights from Research on Teaching                   | 18          |
| Toward a Synthesis                                   | 28          |
| Chapter 3      Ethnographies of Classroom Assessment | 42          |
| Case Study #1                                        | 43          |
| Case Study #2                                        | 58          |
| Case Study #3                                        | 77          |
| Chapter 4      Classroom Assessment Journals         | 137         |
| Analysis of Journals                                 | 138         |
| A Composite Assessment Picture                       | 138         |
| Teachers' Analysis and<br>Criticism of Assessments   | 143         |
| Chapter 5      Summary of Results                    | 151         |

## Chapter 1 Overview of the Report

The research described in this report represents the fourth in a series of studies conducted by the Northwest Regional Educational Laboratory's Center for Performance Assessment intended to provide (a) a clearer understanding of the task demands of classroom assessment and (b) a sense of the assessment training teachers need in order to meet those demands in an effective and efficient manner. The first study in the series was completed in 1982 with NIE funding and provided a very general look at classroom assessment based on discussions with small groups of teachers. The second study, completed in 1983, added considerable detail to our understanding of classroom assessment based on one on one structured interviews with teachers. Study three allowed us to test the generalizability of the results of studies one and two through the development and use of the "Teacher's Self-Analysis of Classroom Assessment"—a comprehensive questionnaire distributed to a stratified sample of teachers across the nation (Stiggins and Bridgeford, in press).

These three studies each contributed significantly to our understanding of the nature and demands of classroom assessment and to our understanding of training teacher needs (Spandel, 1982; Stiggins, 1984; Stiggins, Rubel and Quellmalz, 1985). However, the interview and questionnaire-based research methods used tended to constrain the depth of conclusions we were able to draw. Essentially, we were unable to discern what discrepancies exist, if any, between the assessment methods teachers report using and methods they actually use. The only viable strategies for describing actual classroom assessment practices is to directly observe those practices. Such observations formed the basis of our fourth study of classroom assessment.

Before planning our own observations of classroom assessment environments, however, we conducted a comprehensive review of available research on classroom assessment to take advantage of observational studies previously completed. The results of that review comprise Chapter 2 of the report. These results fed directly into the design of our research.

Specifically, case studies were conducted of 30 classroom assessment environments. Two types of research methods were used to conduct these case studies. Five classrooms were the focus of ethnographic studies of the assessment environment. Researchers served as participant observers for approximately 20 school days to document these key aspects of the methods used to measure student development:

1. What are the purposes served by classroom assessment? Which purposes do these teachers rely on most frequently?
  - a. Assigning grades
  - b. Diagnosing the needs of individual students
  - c. Diagnosing the needs of the class as a group
  - d. Determining student achievement potential (sizing up)
  - e. Grouping for instruction (course placement)
  - f. Selection for advanced or remedial programs

- g. Guidance and educational planning (by the student)
  - h. Feedback to parents (beyond grades, as in conferences)
  - i. Evaluating instructional activities
  - j. Feedback to students (beyond grades)
  - k. Control and motivate students
  - l. Feedback to school managers
  - m. Communicate expectations
  - n. Prepare kids for later assessments
2. What are the various student characteristics measured? Which are measured most often? Most extensively?
- a. Achievement—mastery of materials in a specific subject matter area
    - (1) to determine mastery (criterion referenced)
    - (2) to rank order students (norm referenced)
  - b. Ability to function in higher levels of cognitive operation
  - c. Aptitude—measure the ability to learn
  - d. Social characteristics—as in behavior or interaction
    - (1) of an individual
    - (2) of a group
  - e. Personality characteristics
3. What are the assessment strategies used? Which are used most?
- a. Types of assessments
    - (1) standardized tests
    - (2) text embedded assessment
    - (3) behavioral observation and rating
    - (4) product observation and rating
    - (5) teacher-developed paper and pencil tests
  - b. Key dimensions upon which these might vary
    - (1) planned in advance vs. spontaneous
    - (2) structured vs. informal
    - (3) identified as a test vs. not so identified
    - (4) important vs. unimportant (to teacher, to student)
4. What is the form of the assessment record (or feedback)?
- a. Grade in a gradebook
  - b. Other permanent record
  - c. Written comment (form?)
  - d. Verbal comment
  - e. Non-verbal commentary

The resulting case descriptions comprise Chapter 3.

Thirty-two additional classrooms were studied via teacher journals describing key assessment events. Each journal spanned ten weeks and was conducted in response to this request.

#### JOURNAL ASSIGNMENT

Over the next ten weeks, please keep a journal describing the assessment activities and environment in your classroom. The purposes of this assignment are to (a) raise your level of awareness of how and why you measure student characteristics, and (b) provide me with a profile of how you use your knowledge and skills in the assessment arena.

The journal is to provide a succinct record of the most important assessments and evaluations you conduct. Therefore, to make the assignment manageable, you are to make only one journal entry each week. That entry is to be made at the end of each week and is to describe the single, most important assessment you conducted over the previous week.

Important Note: Please be advised that the most important assessment you conduct need not be a paper and pencil quiz or test—although it may be. The basis of the assessment might be an observation and professional judgment on your part. Further, the assessment need not be a group test. It might focus on an individual student or small group of students. Finally, the assessment need not necessarily be a measure of academic achievement. It might focus on personality, social or affective characteristics. These are all possible candidates for the assessments you describe.

More specifically, select the most important assessment you conducted and describe it in terms of the following points:

1. State the purposes for the assessment and the reason for its importance,
2. Summarize what you wanted to measure (e.g., recall of science facts),
3. Describe how you measured it (e.g., true/false tests, observation, etc.) and why you selected that method,
4. Specify the origin of the assessment (e.g., you developed it, textbook, etc.),
5. Comment on how it worked and how you might revise it in the future.

A summary of journal results is presented in chapter 4.

The purpose of these studies was to define in as much detail as possible all of those ingredients that contribute to an effective classroom assessment environment. Our goals were to identify as many of the factors contributing to the quality of the environment as we could and to describe each factor in terms of a continuum along which any particular classroom might vary. Given such a list of factors we would be in a position, in future research, to generate a profile of the assessment environment in any classroom.

Such a profiling tool would be useful in further study of assessment environments. It will greatly increase the efficiency of our analysis of environments so we can study many more cases, seek a typology of environments, define effective and ineffective environments, uncover the antecedents to ineffective environments and determine those teacher training and resource needs which will allow us to convert ineffective to effective environments.

The list of factors uncovered through observation and journal is presented in the fifth and final chapter.



## References

- Spandel, V. (1982). Classroom application of writing assessment. Portland, OR: Northwest Regional Educational Laboratory.
- Stiggins, R.J. (1984). Evaluating students through classroom observation. Washington, DC: National Education Association.
- Stiggins, R.J. & Bridgeford, N.J. (In press). The ecology of classroom assessment. Journal of Measurement, Vol 22.
- Stiggins, R.J., Rubel, E. & Quellmalz, E. (1985). Measuring thinking skills in the classroom. Unpublished manuscript, Northwest Regional Educational Laboratory, Portland, OR.

## Chapter 2 Insights Into Classroom Assessment

### ABSTRACT

Because the field of educational measurement has tended to concentrate its efforts on research and training in large-scale and paper and pencil assessment methods, teacher-developed classroom assessments have been the focus of minimal research or training. The purposes of this paper are to summarize that which is currently known about measurement in the classroom and to draw implications for future research on testing in the schools. Insights into classroom assessment are gleaned from research on testing, research on teaching (specifically teacher decision making) and classroom ethnographies. The results serve to illustrate a sharp distinction between the scientific assessment that is the goal of the measurement scholar and the practical assessment needs of the teacher. Many unanswered questions about the classroom assessment environment are identified and reasons are presented for generating answers as soon as possible. Teacher and administrator training priorities are also considered in light of the research review.

## THE MEASUREMENT PARADIGM

In any field of inquiry, Kuhn (1970) has told us, we can expect scholars to adopt a set of conventions of research design and concept development to maximize the efficiency of their communication and the productivity of their research efforts. Those conventions define the dominant paradigm for that field of inquiry:

"a paradigm is an implicit, unvoiced and pervasive commitment by a community of scholars to a conceptual framework. In a mature science, only one paradigm can be dominant at a time. It is shared by that community and serves to define proper ways of asking questions...[and of defining] those common "puzzles" that are defined as the tasks of research [in that field]." (Shulman, 1985, p. 8)

What paradigm guides scholarship in educational measurement? Available evidence suggests that we tend to regard measurement in education as a process of documenting the student achievement using collections of standard paper and pencil test items for purposes of public accountability. Evidence of the dominance of this conceptualization can be found in measurement textbooks, research reported in scholarly journals, and in published standards of accepted professional practice in measurement.

Consider, for instance, the message conveyed to teachers and graduate students in these opening sentences from an introductory measurement textbook by Mehrens and Lehmann (1984):

Educators have always been concerned with measuring and evaluating the progress of their students. As the goals of education have become more complex and with the increasing demand by all parts of our citizenry—pupils, parents, taxpayers and other decision makers—for accountability on the part of educators, these tasks of measurement and evaluation have become more difficult. (p. v)

These authors go on to instruct that a test is defined as "a standard set of questions to be answered...[from which] we obtain a measure (that is, a numerical value) of a characteristic of that person." (p. 4) In fact, Mehrens and Lehmann go on to describe a variety of other forms of and purposes for assessment. But the dominant form of assessment covered in their instruction to teachers relates directly to the quantification of achievement via collecting of various types of standard paper and pencil test items.

Consider also the message about educational measurement contained in these opening sentences of another introductory textbook by Ebel (1979).

There is a paradox in educational measurement today. While assessments of achievement and competence are being more urgently called for and more widely employed than ever before, at the same time tests are being more sharply criticized and strongly opposed. Perhaps these apparent inconsistencies are in fact wholly consistent. If demands for more measurement of educational outcomes arise out of dissatisfaction with how much is being learned, those responsible for the teaching are likely to feel threatened by possible exposure of educational shortcomings. (p. 1)

The message that assessment should serve an accountability purpose comes through loud and clear. Subsequent instruction on test development teaches students about tests defined as follows: "The most commonly used types of tests are the essay type, the objective, the mathematical problem type and the oral examination type." (p. 56) For all practical purposes, the entire text deals with the design, construction, evaluation and use of tests comprised of standard paper and pencil test items.

This pattern of dominance of one form of assessment is seen repeatedly in the textbooks that introduce educators to the assessment process. Please note that our only purpose is to establish the clear dominance of a certain point of view. We do not claim that these texts address only this one form of assessment. To their credit, text authors expand the paradigm to include tests from two key sources—teachers and professional test developers. Other evidence, however, suggests that the dominant paradigm is much narrower in scope.

For instance, an examination of research and scholarly writings in the field reveals a clear narrowing of the scope to a focus on the standardized tests developed for and used in large-scale testing programs (i.e., for accountability purposes). For instance, until very recently, nearly all major studies of testing in the schools focussed on the role of standardized tests (Goslin, 1967; Lortie, 1975; Airasian, et al, 1977; Stetz and Beck, 1979; Rudman, et al, 1980; Salmon-Cox, 1981; Sproul and Zubrow, 1982; Kellaghan, et al, 1982; Fyans, 1985; and Tollefson, et al, 1985). Further, in a recent special issue of the Journal of Educational Measurement (Burstein, 1983) on the "state of the art integrating testing and instruction," the editor introduced the issue as follows:

Linking testing and instruction is a fundamental and enduring concern in educational practice...Fundamental questions about how well achievement test items reflect both student knowledge and the content of instruction are clearly at the heart of the matter...[yet] The contributors [to this special issue] were asked to limit their conception of achievement testing to include standardized achievement tests, curriculum embedded or locally developed domain-referenced and proficiency tests, and state assessments. Thus, teacher made tests...were systematically excluded. (p. 99, emphasis added)

A review of the four most recent volumes of that same journal conducted by Stiggins and Bridgeford (1985) revealed that nearly all reports dealing with the measurement of achievement dealt with topics relevant to the use of objective paper and pencil tests and a vast majority of those focused on topics most relevant to standardized--not teacher-made--tests.

But the final and perhaps most telling piece of evidence that standardized tests dominate current thought about educational measurement is the fact that the only written standards on acceptable testing practice are the Revised Standards for Educational and Psychological Tests (APA, 1984) which detail the ethical responsibilities of publishers of standardized paper and pencil tests. No such standards exist for teacher-made tests. No such standards exist for tests relying on other than paper and pencil test items.

### Utility of the Dominant Paradigm

The extreme dominance of this conceptualization of educational measurement over the past four decades is testimony to its utility. As Coffman (1983) and Calfee and Drum (1976) point out, it has afforded education an image of scientific precision and ultimately has fostered a tradition of scientific inquiry in educational research and psychometric theory. Politically, it has given educational measurement a visible role in documenting the effectiveness of schools in our society. The coin of the realm in determining the value of schools is clearly the standardized test score.

However, it is the premise of this paper that the cost of providing only outcome measures of school effectiveness research has been very high, particularly in terms of the contribution of educational measurement to research and development in teaching and learning. As the research summarized below will reveal, the kind of measurement referenced under the dominant paradigm represents only a small fraction of the assessment that impacts the quality of schooling and impact of schools on students. Unfortunately, however, due to our narrow vision of measurement research, we know little about the nature, role, or quality of the remaining assessment. One price we have paid for our concentration on large-scale standardized paper and pencil tests of achievement is that we have neglected research on teacher-developed classroom assessment.

This neglect has led us to a crucial conflict for the field of educational measurement as we proceed through the decade of school improvement. One potential role for measurement is as a process variable in school improvement. That is, measurement could be considered one of those skills like classroom management that can be the focus of research and training. Teachers might be trained to measure better so classroom assessment is improved and the quality of overall instruction is improved.

Or, measurement might serve as an outcome variable, providing the index by which we measure the success of other instructional improvement efforts. The dominant paradigm casts measurement in this role--an accountability role, a passive role. So we concentrate our resources on ensuring the quality of standardized outcome measures.



We are unable to do otherwise, because we know little about classroom assessments nor how they might be improved. To illustrate, Clark and Peterson (1985) after reviewing all available research on the relationship between teacher behavior and student achievement, are able to draw conclusions about only one small facet of effective teacher-developed assessment practices--the management of oral questioning during recitation or discussion. It is equally troubling to have those most deeply immersed in the vast and growing body of research on teaching conclude, "In general, the kinds of tests we use are inconsistent with, and in many cases irrelevant to, the realities of teaching." (Shulman, 1980, emphasis in original).

Lazar-Morris, et al (1980) conclude their comprehensive review of research on testing in the schools as follows:

In-class assessments made by individual teachers have yet to be examined in depth. How these and other assessments are united with teacher instructional decision-making processes and how they affect classroom organization and time allocation to other objectives are areas that should be explored. Teachers place greater reliance on, and have more confidence in, the results of their own judgments of student performance, but little is known about [these] kinds of activities. (p. 24-25)

The unfortunate consequence of this neglect is that the measurement field missed a golden opportunity to play a key role in school improvement through better classroom assessment.

#### Focus of the Review

The purpose of this paper is to summarize what is currently known about classroom assessment in the hope that it will foster further movement toward a broadening of the dominant view of educational measurement and thus expand our measurement's role in the process of school improvement. Three bodies of research are reviewed. First, we explore for insights into classroom assessment arising out of the body of knowledge on testing in the schools. Then we turn to the growing collection of studies on teacher decision making for insights as to the nature and quality of assessments used in decision making. And finally, we will turn to available ethnographies of classroom assessment for material descriptive of classroom assessment. In each case, we explore their implications for teacher training in educational measurement. In conclusion, we will look beyond available information to a series of as yet unanswered questions about classroom assessment, and will explore implications for future measurement research.



## INSIGHTS FROM RESEARCH ON TESTING

Although research on testing in the schools has tended to concentrate on the role of standardized tests, a few of these studies provide insights into the nature of the classroom assessment environment. These studies are reviewed in this section, focusing specifically on what they tell us about the relative importance of various types of assessment, assessment processes in the classroom, the quality of classroom assessments, and prevailing teacher and student attitudes about assessment.

### Relative Importance of Assessment Types

Studies of testing in the schools tells us that teachers rely on their own assessments as the primary source of information on student achievement. The following report of research by Morine-Dershimer (1979) and Joyce (1979a,b) depicts the importance teachers attach to their own measurements:

Morine-Dershimer and Joyce observed the reactions of the teachers when a set of domain-referenced diagnostic tests that the state had mandated was returned for each of the classroom teachers' use. Performance of each pupil was keyed to each objective and, if pupils were low, the printout specified what kind of curriculum materials could be used to remediate the deficiency. The investigators waited until two weeks after the tests had come back to interview the teachers because they wanted to study how teachers' conceptions of their pupils had changed since the beginning of the year, especially after this marvelous new set of information had arrived. It turned out, however, that not a single one of the ten teachers had looked at those test results. They simply did not find them useful. They were convinced that they already knew more about their students than any one of those tests could possibly detect. Most of the teachers did not believe the tests were of any value at all. (Shulman, 1980, p. 68)

Studies conducted at the Center for the Study of Evaluation (Herman and Dorr-Bremme, 1982 and Yeh, 1980) suggest that, depending on grade level, a third to three-quarters of assessments used in classrooms are teacher developed. Other studies reveal quite clearly that those assessments include far more than collections of paper and pencil test items. For example, Herman and Dorr-Bremme (1982) report, "nearly every survey respondent reported that 'my own observations and students' classwork' was a crucial or important source of information." In another study, Salmon-Cox (1981) concludes, "overwhelmingly, we found that teachers, when talking of how they assess their students, most frequently mention 'observation'. Clearly this favored teacher technique is quite different from the kind of information provided by standardized tests." And Kellaghan, Madaus and Airasian (1982) point out "standardized test information appears to represent an auxiliary or secondary criterion in [instructional] judgment, since teachers were nearly unanimous in stating that the most commonly reported grouping criteria were the teachers' own observations and tests."

In fact, Stiggins and Bridgeford (1985) surveyed 228 teachers across the nation to determine preferential use of various assessment types. Each teacher was asked to distribute 100 points across four types of assessment (teacher-made objective tests, standardized tests, structured performance (behavioral) assessments, and spontaneous observations) to convey the relative importance of each type. Although there were some differences as a function for assessment purpose, grade level and subject, teachers assigned an average of 34 points to their own objective tests, 26 points to structured performance assessments (preplanned observation and judgment strategies), 21 points to spontaneous observations and judgments and 19 points to published tests (including curriculum embedded and standardized tests). Salmon-Cox (1981) found that, of 87 high school teachers she interviewed, 44% reported using their own tests for evaluating students, while 30% used interaction, 21% relied on homework, 6% used observation and none reported using standardized tests.

These data suggest that, if our goal is to ensure quality assessment in the classroom, we need to focus research on understanding the quality of assessments relied on by teachers. We have not done so. Heartel, et al (1984) reviewed recent summaries of available research on testing and concluded as follows:

From the results summarized, it can be seen that the majority of information currently available on perceptions of testing in schools focuses on teachers' attitudes toward norm-referenced, standardized testing. Far less information is available on students' perceptions of and attitudes toward the tests they take throughout their academic lives. (p. 7)

Airasian (1984a) echoes this view:

If the present state of knowledge makes it difficult to describe the full complexity of classroom assessment, it does not prevent us from making some generalizations which are useful in focusing concern and providing guideposts for discussion. In spite of a great deal of work which has been done on classroom assessment in recent years, distinctions among the varied purposes of classroom assessment have not always been identified or articulated. (p. 8)

#### The Nature of Assessment Processes

Despite the neglect, measurement research has been able to provide some insight into the nature of classroom assessment. We know, for instance, that classroom assessment environments are designed and constructed by teachers who have had little formal training in assessment (Ward, 1982; Coffman, 1983). Many have had no formal course work and most no inservice training in the subject. Further, the analysis of textbooks discussed in the introductory section of this paper noting the emphasis on objective testing issues suggests that the training conducted was narrow in focus. Analyses by Stinnet (1969), Woeller (1979) and Burdin (1982), reveal no requirements that teachers be trained in testing to be certified.

In spite of an apparent lack of formal training, however, teachers are able to employ a wide range of measurement strategies. This point was clearly established in the previous section. However, Stiggins and Bridgeford (1985) provide further amplification of teacher use of assessment methods by exploring the importance teachers attach to different forms of assessment as purpose, grade level and subject matter varies. They conclude that this sample group of teachers are quite consistent in the assessment methods they use across purposes. As purpose varies from diagnosis, to grouping to grading, etc. the relative importance teachers attach to different forms of assessment remains quite constant. Further, of the 228 teachers surveyed, less than 5% reported taking any action toward revising their current testing patterns.

However, there is evidence of fundamental differences in the nature of assessment as grade increases. For instance, Herman and Dorr-Bremme (1982) report that 75% of the tests used by the over 350 high school teachers they surveyed were teacher developed, while the over 400 elementary teachers relied more heavily on curriculum embedded tests (tests included in text materials). Stiggins and Bridgeford (1985) report that when assigning grades, relative importance of different types of assessment changes with grade level: "As grade level increases, the weight given to objective tests and structured performance assessment goes up, while that given to published tests and spontaneous observations and judgments goes down." (p. 10)

Stiggins and Bridgeford (1985) also report differences in the relative importance of different assessment processes as a function of school subject. Math and science teachers, as expected, tend to rely most heavily on paper and pencil objective tests, while teachers focusing on communication skills (writing and speaking) rely more heavily on structured observations and professional judgments.

Airasian (1984b) summarizes other dimensions of classroom assessment which have been researched.

- o There appear to be two sets of characteristics measured in classrooms--scholastic variables and social variables (Airasian, Kellaghan & Madaus, 1977; Herbert, 1974; Pedulla, Airasian and Madaus, 1980).
- o The relative importance assigned to these two factors varies with grade level, with social factors seen as more important in elementary school (Salmon-Cox, 1981).
- o Teachers "size up" students as individuals and group very quickly and these initial estimates remain quite stable (Rist, 1970; Airasian, Kellaghan & Madaus, 1977).
- o Students appear sensitive to these early teacher assessments, learn their positions in the "pecking order" of the class and respond accordingly (Morrison and McIntyre, 1969; Rist, 1970).

- o Teachers interact differently with students they perceive to be of high or low ability (Brophy and Good, 1974).
- o Teachers can accurately predict student test performance and thus use standardized test results to corroborate their own judgments (Kellaghan, Madaus and Airasian, 1982).

Three in-depth studies have been conducted of the characteristics of teacher-made tests. Fleming and Chambers (1983) report the results of an analysis of well over 300 teacher-made paper and pencil tests, and Stiggins and Bridgeford (1985) report the characteristics of teacher-developed structured performance assessments. Fleming and Chambers drew these conclusions about teachers' paper and pencil tests:

First, teachers use short-answer questions most frequently in their test making. Second, teachers, even English teachers, generally avoid essay questions, which represent slightly more than 1 percent of all test items reviewed. Third, teachers use more matching items than multiple-choice or true-false items. Fourth, teachers devise more test questions to sample knowledge of facts than any of the other behavioral categories studied. Fifth, when categories related to knowledge of terms, knowledge of facts, and knowledge of rules and principles are combined, almost 80 percent of the test questions reviewed focus on these areas.

Sixth, teachers develop few questions to test behaviors that can be classified as ability to make applications. Seventh, comparison across school levels shows that junior high school teachers use more questions to tap knowledge of terms, knowledge of facts, and knowledge of rules and principles than elementary or senior high school teachers do. Almost 94 percent of their questions address knowledge categories, contrasted with 69 percent of the senior high school teachers' questions and 69 percent of the elementary school teachers' questions. (p. 32)

In another study, Carter (1984) studied the test development skills of 310 high school teachers and reports that teachers had great difficulty recognizing items written to measure specific skills—especially higher order thinking skills. She also reports that teachers learned to write original items at higher skill levels very slowly and felt insecure about their test making capabilities.

In their research on teachers use of performance assessment, Stiggins and Bridgeford (1985) explore those assessments in which "students are called upon to apply the skills and knowledge they have learned...[through the] completion of a specified task in the context of a real or simulated assessment exercise...[in which] the process or product completed by the examinee is



observed and rated by the teacher" (p. 5). Over three quarters of the 228 teachers surveyed reported using this form of assessment and described them as follows. They tend to be:

- o equally divided between evaluations of processes (students performing as in speaking) and products created by students,
- o scored both holistically and analytically but result in a single grade being assigned,
- o scored by the teacher rather than the student or a colleague,
- o interpreted in criterion referenced terms—in terms of a pre-established standard,
- o public and preannounced rather than unobtrusive assessments, and
- o are often used with little attention to assessment quality.

#### The Quality of Classroom Assessments

Research on testing in the schools has provided very little information concerning the quality of teacher-developed assessments. And that which is available is quite narrow in scope. For instance, we can infer that some teacher-developed assessments have validity, since teacher-based assessments allow some teachers accurately to predict student performance on standardized achievement tests (Kellaghan, et al, 1982; and Fyans, 1985). Further, teachers often feel their tests are valid (Farr and Griffin, 1973).

However, there are some indications of potential quality problems. For instance, Fleming and Chambers (1983) and Carter (1984) cite a need for teachers to write better test items—particularly items that are less ambiguous and take students beyond the simple recall of facts and information. In addition, Stiggins and Bridgeford (1985) report inattention by teachers to those procedures that are likely to promote valid and reliable performance assessment, such as clearly articulating and communicating scoring criteria, defining acceptable levels of performance, repeating observations, keeping written records and checking judgments against other data such as test scores. Further, they found that attention to these quality control steps increased as grade level increased.

There is some preliminary evidence that teacher-developed tests are very short i.e. contain a minimal number of items (Fleming and Chambers, 1983). Test length can impact test reliability. It may also be that some teachers are insensitive to the potential problems with their assessments. A lack of knowledge of what might go wrong precludes (a) test design to prevent problems and (b) investigation to see if potential problems in fact arose. Time pressures may also preclude careful test design.



Gullickson's (1982 and 1985) studies of South Dakota teachers' testing strategies further reinforce a lack of quality control strategies. For example, few of the teachers surveyed compute summary statistics needed to evaluate test performance; most limit test questions to short answer and matching both of which test lower cognitive levels; few teachers take time to improve their tests; and teachers usually reuse items without careful item analysis. Overall, Gullickson concludes that teachers do not know how to evaluate their test items; take necessary steps to improve quality; or to accurately set criterion levels for student performance. Further, they do not value statistical analysis of test items as a helpful strategy in the classroom (Gullickson, 1984a, 6).

Research on teaching has also provided some information on the reliability of teachers' assessments. That research is discussed in a later section.

### Attitudes about Assessment

Although most research on testing in the schools has focused on attitudes toward standardized tests, a few studies allow us to draw some conclusions about both teacher and student attitudes about classroom assessment. Some of those attitudes are reflected in the patterns of test use among teachers. Teachers value assessments that provide information relevant to the decisions they may face. Salmon-Cox (1980) interviewed 35 elementary school teachers who articulate some of those values. Teachers judge students even in the absence of formally communicated information. They give social and background characteristics greater emphasis than ability in classroom assessment. And observation of students is the most frequently used mode of assessment for these teachers. She would later conclude that "teacher preference, in effect, is for continuous movies, with sound, while a test score or even a profile of scores, is more akin to a black and white photograph." (Salmon-Cox, 1981)

Stiggins and Bridgeford (1985) explored teacher attitudes about testing by asking teachers to indicate their concerns about various types of assessment. By far the most frequently expressed concern was uncertainty about how to improve test quality and to manage the assessment environment. Although teachers in this sample were not in the midst of changing their assessment methods and were generally comfortable with their current procedures, they consistently noted that they were interested in suggestions for improvement.

Student attitudes about standardized tests have been studied to some extent, but again until very recently, researchers have expressed little interest in exploring student perceptions of teacher-developed assessments. One study of standardized tests included some student-directed questions about classroom assessment. That was conducted by Stetz and Beck (1979) and reports that students are more concerned about teacher-made than standardized tests.

Most think teacher-made tests are harder and twice as many get nervous before a teacher-made test. Heartel's, et al (1984) analysis of questionnaire responses of over 600 high school students provides a more comprehensive perspective. They conclude that

Students conceive of tests as limited to formal, paper-and-pencil assessments, usually asking objective questions, and quite separate from ongoing instruction. The purpose of testing is primarily to assign marks and grades. While students consider tests important and are willing to work to earn high scores, they see tests as requiring mostly memorization, perhaps to the detriment of other types of learning.

Students understand that there should be more to schooling outcomes than answering multiple-choice questions; over half recognize that many important ideas are not tested at all. Nonetheless, while students may feel that they know more than their test scores show, they are most comfortable with the familiar true-false and objective types of items; they dislike testing formats that require more extensive response. (p. 29)

## INSIGHTS FROM RESEARCH ON TEACHING

Research on teaching has focused on relationships between teacher behavior and student outcomes, relationships between student behavior and learning, classroom processes, and the cognitions of both students and teachers as instruction proceeds (Shulman, 1985). The purposes of this research have been to explicate those models of teaching that help us better understand the teaching and learning process so teacher training experiences can be designed to maximize teacher effectiveness.

Although contemporary research on teaching has not focused specifically on effective classroom assessment, this body of research does provide some interesting and useful insights to supplement what research on testing has shown us about the nature of the classroom assessment environment. Recent summaries of research on teaching compiled by Shavelson and Stern (1981), Clark and Peterson (1985) and Shulman (1985) instruct us in two specific areas. First, these summaries provide a window into teacher decision making processes, allowing us to see the complexity of the teacher's classroom assessment task. Second, we can use this window to explore the nature and role of assessment before, during and after instruction. From this vantage point, we can see the great challenges teachers face in accurately assessing student characteristics.

As we look at measurement in the classroom and in research on teaching, we will point out implications for measurement research, development and training.

### The Complexity of Classroom Assessment

Research on teaching tells us that assessment is unquestionably one of the most complex and important tasks faced by teachers. Each investigation of the teaching process arises out of a model or conceptualization of teaching and learning. Literally every model of effective teaching requires that teachers base their instructional decisions on some knowledge of student characteristics. We begin to comprehend the complexity of classroom assessment as we explore the range and frequency of the decisions teachers make and the plethora of student characteristics they must consider in making those decisions.

Investigations of classroom practices have tended to focus on three major types of decisions, each placing significantly different measurement demands on teachers. These have been labeled preinstruction (preactive, planning) decisions, interactive decisions (made during instruction), and post instructional decisions. Shavelson and Stern (1981) summarize 30 studies of teacher decision making, identifying the type of instructional decision teachers faced in each and the salient cues those teachers considered in making those decisions. Salient cues represent specific student characteristics considered. Sixty-six such cues were listed across all studies. If we classify these cues as representing academic, social, or personality student characteristics, and employ the preinstructional, interactive and postinstructional categories to group decisions, and cross the

two classes as in Table 1, what results is a frequency count reflecting (1) the extent to which teachers must be able to measure more than academic achievement and (2) how factors considered vary as a function of the nature of the decision.

**TABLE 1**  
Frequency of student characteristics reported in  
research by type and decision context studies  
(adapted from Shavelson and Stern, 1981)

| Characteristic | Decision Context |                    |                   | Total |
|----------------|------------------|--------------------|-------------------|-------|
|                | Planning         | During Instruction | After Instruction |       |
| Academic       | 15               | 11                 | 2                 | 28    |
| Social         | 5                | 15                 | 3                 | 23    |
| Personal       | 6                | 4                  | 5                 | 15    |
| Total          | 26               | 30                 | 10                | 66    |

When faced with planning decisions, greatest reliance was placed on academic and ability variables as antecedent variables. Decisions made during instruction had antecedents in social interaction along with academics while decisions rendered after considered a variety of salient cues. Clearly teachers must measure more than achievement.

#### A Closer Look

Available research also allows a closer look at classroom assessment by concentrating on those studies focusing specifically on teachers' decision making during instruction (interactive decisions). Some of the other student characteristics considered by teachers according to this research are listed below. These too give testimony to the complexity of classroom assessment.

##### Social Characteristics

Disruptiveness  
Work habits  
Classroom behavior  
Consideration of others  
Group Mood  
Participation  
Involvement  
History of deviant behavior  
Behavior  
Peer relations

##### Personal Characteristics

Motivation  
Self-confidence  
Sex  
Determination  
Openness to new ideas  
Sense of humor  
Activity level  
Attentiveness  
Attitudes  
Family background  
Health  
Personality  
Cooperativeness  
Maturity



But perhaps the most important point arising from the above lists from a measurement perspective is that these factors are not just considered in managing disruptive behavior. They play a role in planning instruction, managing interactive exchanges and in evaluative judgments about students. Further, when teachers gather information about these factors, they have no published standardized tests to rely on. They are left to their own devices, with little support or training. More about that later.

Most models of interactive decision making (during instruction) have the teacher observing some form of student behavior or performance and comparing it to a standard to see if the sample is within tolerance (Yinger, 1977; Shavelson & Stern, 1981; Clark and Peterson, 1985). These decisions occur frequently. For instance, Clark and Peterson (1985) synthesized six studies and concluded that teachers make an interactive decision on the average of every two minutes. To understand the measurement implications of this pace and the importance for learners, consider the fact that half of these decisions have antecedent thoughts on the part of the teacher based on concerns about the learner, including comparisons of behavior, etc. with expectations or standards (Clark and Peterson, 1985; and Marland, 1977). In this context, the teacher must either assess very rapidly with validity and reliability or rely on an existing reservoir of valid and reliable information. Surely this is an assessment demand unparalleled in other professions.

As if interactive decision making were not complex enough, we must also consider the pre- and post-instructional decisions teachers face. Planning is a complex enterprise for teachers. In this case, however, the research holds a surprise. Work conducted by Zaharik (1975), Yinger (1977), and others (summarized by Clark and Peterson, 1985) reveals that teachers do not tend to focus on assessments of student characteristics when they plan. Nor do they focus on goals and objectives. The instructional activity is most often the planning unit. Teachers focus on activities and content—what they will do and cover. Shulman (1985) explores the implications of these data for measurement and evaluation:

For years, those of us in educational research, especially in evaluation and measurement, have been insisting that teachers learn to think straight educationally. By that we mean they have to learn to think of outcomes stated in terms of behavioral objectives. However, if generations of practitioners do not think in such a way, an alternative consideration might be that there is something adaptive in focusing instead on activities and content covered.

Teachers appear not to evaluate their day-to-day activity in terms of general assessments of achieved outcomes, but rather attend to variations in student involvement. When we ask teachers, "What did you achieve today?" they are inclined to say, "Well, we covered three more pages of math, and the kids were really involved." We then become critical and berate teachers for not thinking in terms of objectives—which ones they achieved and which not. I believe we have to treat the teachers' observations as data rather than as sources for blame. That is how teachers evaluate what they do. When they plan their instruction, they plan for such things as grouping, pacing, and involvement. (p. 70)



This tendency on the part of teachers to focus on accomplishments may also have implications for the evaluation of students. Although no research has addressed this issue specifically, it may be that in grading, teachers rely more heavily on what the student has done—tasks completed—and less heavily on the quality of work or what was learned as a result of the experience. More about this later.

When teachers do focus on student characteristics during planning, it is often only early in the year and it is done very quickly and efficiently. Calderhead (1983) and Salmon-Cox (1980) point out that experienced teachers become very proficient at using available information to understand their new class almost before it arrives in the classroom. However, these rapid conclusions based on scant data can have detrimental effects. Peterson and Barger (in press) suggest that teachers may become fixed on initial impressions and use subsequent data to maintain a consistent picture of the student. That is, it may be difficult for a student to break down an initial incorrect impression—one that was based on unreliable, inaccurate information.

### Dealing with Complexity

Given this impressive array of decision contexts and student characteristics, how do teachers proceed? The research on teaching provides some clues. For instance, Simon (1957) suggests that, when faced with an overload of information to process, teachers simplify their view of reality, thus creating a manageable task. Teachers find algorithms and heuristics that allow them to process and store information parsimoniously, often as a reflex action.

For example, teachers may reduce the number of assessments to be conducted by tapping groups rather than individual data. Dahlof and Lundgren (1970) found teachers identifying a "steering group"—a select subset of students in the class whom they could check for reliable information on whether to repeat instruction or proceed to the next topic. Jackson (1968) illustrates the nature of some of the group cues teachers zero in on to evaluate whether instruction is working:

Oh, look at their faces...they look alert, they look interested, they look questioning. They look like they're anxious to learn more about it...And other times you know you haven't done a good job when they look blah or disinterested or (show an) I don't care attitude.

A theatrical sense is something that you can't learn, but a good actor can sense his audience. He knows when a performance is going well or not going well, simply by the feeling in the air. And it's that way in the classroom. You can feel when the kids are resistant. (p. 122)

As mentioned earlier, it appears that teachers gather information very quickly and form it into impressions of student ability very early in the school year (Calderhead, 1983). And once those judgments are made assessment

of ability ceases leaving those impressions solidly in place and allowing the teacher to move on to other aspects of the complex assessment task. There is also evidence that teachers turn to characteristics that are most easily measured such as social behavior and task completion which often become as crucial as achievement in classroom assessment (Weiner and Kukla, 1974). And, as research on testing showed us in the previous section, when teachers measure achievement, they focus on those levels of achievement most efficiently measured, such as recall of facts (Fleming and Chambers, 1983).

How successful are teachers at dealing with the complexity of the assessment environment? As both Rudman et al. (1980) and Shavelson and Stern (1981) clearly indicate, teachers' decision making strategies are the focus of an increasing number of studies. Shavelson and Stern's (1981) literature review on teachers' pedagogical judgments, decisions and behaviors provides a thorough examination of the research being currently undertaken in this area of inquiry. Outcomes from this review plus a number of additional studies, briefly summarized here, indicate the diversity of information on teacher decision making--and judgment processes--issues at the heart of the instructional and evaluative role of the teacher.

Shavelson, Caldwell and Izu (1977), for example, conducted a laboratory-based experiment to determine whether teachers reconsider their initial estimates of student abilities when presented with new information and if they considered the reliability of information when making judgments. The outcomes of this study indicated that teachers revised estimates of students when presented with new and differing information and that teachers showed appropriate sensitivity to the reliability of information sources. Similarly, a number of other studies noted in Rudman, et al. (1980) comment equally favorably on teachers' ability to correctly judge students' ability and successfully estimate students' performance on standardized tests.

Other studies, however, recount numerous problems in teachers' judgments. For example, Brophy and Good (1970) concluded that teachers' expectations of students clearly correlated with differential patterns of interaction between teacher and student. After coding the behavior of teachers and students, the researchers note that teachers consistently favored the highs over the lows in demanding and reinforcing quality performance. Students perceived with high expectations were more frequently praised when correct and less frequently criticized when incorrect or unable to respond. A later study of this same nature Good and Brophy (1978) verified similar outcomes and also indicated that teachers were totally unaware of their different levels of interchange with students who they judged to have different abilities. Weinshank (1980), in a multi-phased investigation of the clinical problem-solving skills of reading and learning disability specialists, discovered surprisingly low diagnostic and remedial reliability among specialists. Wienshank noted that the mean agreement between any two clinicians on a given case (0.08), was no more than would have occurred if it were a chance occurrence. Surprisingly, clinicians were almost as unlikely to agree with themselves when presented with a replicate case, as they were with other clinicians. Rudman, et al (1980) also recount similar investigations that substantiate low reliability in diagnostic evaluations of students by reading specialists. The studies by

Weinshank and others provide some of the most conclusive evidence of problems with teachers' ability to diagnose students' skills and, hence, with the accuracy of their judgments. As Shavelson and Stern (1981) point out, teachers' assessments of behavior are the critical factor in determining instructional decisions and "Teachers' judgments about students, for example, and not the original information about students, appear to be the basis for decision making" (p. 475). Research by Gil and Freeman (1980), also confirms the inadequacy of teachers' judgment procedures in natural as well as laboratory settings. After observing and interviewing ten teachers, the researcher concluded that teachers clearly lacked "information-processing strategies to make complete, specific diagnoses." (Shavelson & Stern, 1981, p. 477)

Although most of these foregoing studies on decision making (with the exception of Gil's work) were conducted in laboratory settings, the gradual incorporation of naturalistic observation in classrooms has not yet provided more encouraging results on teachers' assessment accuracy. Whitmer (1983), in studying the judgment process of five elementary school teachers during marking, found an emphasis on completion of tasks rather than quality of outcomes in teachers' procedures. "The factor of completion, or the filling in of columns across the teachers' record book, appeared to carry a heavier weight than the quality of completed work." (p. 26) Teachers failed to consider the level of difficulty of tasks in grading students or to weight various tasks or assignments in deriving summative grades. Teachers' emphasis on expediency and simplicity in grading, often at the expense of careful discrimination about student work, further reinforces the assumption that teachers' practices in evaluating students lack rigor and possibly appropriateness.

Again, these findings from research on teaching are preliminary, arising from a few studies which focus on a narrow set of subjects and grades. Very few teachers have been studied with respect to the reliability of their assessments and judgments. But there exists at least the danger of fallible assessment information--especially in the assessment and decision-making world described in this research. And in this case, the impact will be on the achievement and academic and personal self-concept of an individual learner.

### Classroom Ethnographies

Ethnographies of the schools have focused particularly on interactions in multicultural classrooms. Many of these studies have investigated teacher decision making, particularly the role of ethnicity and culture in the evaluation of students' informal classroom behavior and their performance on assigned tasks. More recent work has also considered classrooms in which the children and the teacher all have similar cultural backgrounds, but in which the interactional messages from teacher still present difficult decoding challenges to the students.

Although ethnographers have not addressed assessment in the terms familiar to and used by assessment researchers, their studies of classroom interaction



are directly relevant to assessment research. They have, for example, studied "teacher assessment of the intellectual competence of children on the basis of social performance", including

(1) the cues teachers employ to make judgments of competence--e.g., how children talk, listen, sit, respond to procedural instructions...; (2) the relative differentiation of the teacher's typology of children in the class--the range of "taxons" or dimensions of contrasts in the teacher's cognitive map of the kind of students in the class; and (3) the relative stability of the teacher's typology over time. (Erickson, 1977, p. 64)

In this section, we explore some of the results of these studies and their implications for decision assessment.

Crosscultural Studies. Pioneering work in this area was undertaken by Labov (1970, 1972) and focused first on oral standardized testing situations. He found that most black children performed very poorly on such tests, tracing the problem to cultural attitudes toward the test situation. Ethnographers argued that the children's extremely verbal behavior in unmonitored or less structured classroom situations belied these negative evaluations of language and cognitive ability and suggested that the testing situation itself was an uncontrolled factor in standardized testing.

These ethnographic studies have long since led to study of less formal classroom testing techniques and general classroom interaction. In many cases researchers have discovered explanations for school failure in mismatches between culturally determined behavior patterns that the children bring to the classroom from their homes and communities and those prevailing in the classroom. In monocultural classrooms as well they have found that failure to understand the teacher's norms and codes of interaction have been misdiagnosed as failure to comprehend instructional content.

Classroom ethnographies usually focus on verbal interaction between teachers and students. Ethnographers of schooling have frequently studied question and answer sequences because classroom teachers often test children by asking questions. This strategy demands that individual children perform for their teacher and the class, demonstrating their knowledge as succinctly or as expansively as possible, depending on the particular exercise. This type of interaction requires that the children be willing--in fact eager--to display their talents before others and that they be prepared to risk failure before their peers as an acceptable route to learning.

Parental standards for polite behavior serve as children's models for interactional success. These standards may or may not be congruent with those expected by and promulgated by the schools. As Heath (1982 and 1983) has shown, incongruities may be recognized by the children, who adapt if they can, at the danger of acting in an inappropriate fashion when at home. Adaptation may prove too difficult, if the expected behavior is embarrassing to them personally or damaging to their relationships with peers. Or children may simply fail to recognize or understand the behavior expected of them and continue to adhere to their own cultural patterns.

For instance, known-answer questioning in the classroom has proven unsuccessful with American Indian children. In fact, teachers and ethnographic researchers working with many different tribes report that any form of classroom interaction that singles out one pupil to act as performer seems doomed to failure (Philippe 1972, DuMont and Wax 1969). These children are reluctant to read aloud and, if forced, will be accompanied by the whispered recitation of their classmates, who supply the text, should the speaker stumble over an unknown word. Rather than vying with one another in assessment contexts, Indian students more often tend to function solidly as a group, aiding fellow students by providing difficult answers or distracting the teacher's attention from an individual in danger of reprimand.

Oral speaking style has also been the subject of considerable study. Cooley (1979) studied American Indian students enrolled in a freshman college speech communications class and uncovered an approach to public discourse that was very much at odds with the instructor's model of persuasive speech. The instructor regularly failed Indian students because their speeches were not organized around a single central topic, nor did they make arguments substantiating their point of view or draw connections between apparently unrelated remarks. Rather, the Indian students presented a series of arguments relating to their topic, some substantiating and some contradicting each other. The listener was left to draw connections among them; no point of view was articulated by the student speaker. Performance criteria were either unclear for students or out of reach.

Cooley extended his inquiry to the Indian community, attending public meetings and pow-wows. There he discovered the models for the students' classroom speeches. Tribal leaders did not interpret information for their audience; they presented all the known facts and deferred decision to the group after all the presentors had been heard. He concluded that

It is the role of the listener to put that information together and to arrive at a conclusion about its worth or about how it applies to the subject at hand. Any overt marking of the relationship between topics, whether by the use of transition devices or of cohesive devices, could be construed as an attempt to lead the audience toward a decision, and that would be improper (p. 557).

A study of a very different population yielded analogous findings. Michaels and Cook-Gumperz (1979) studied oral storytelling in a first grade classroom and found that white children attempted to construct their stories using a "topic-centered" style in which all description is subordinated to and directly related to a single, central point. However, black students tended to construct stories that were "topic-chained". Topic-chained stories are constructed of a series of comments, each with its own topic; the connections between them are not made explicit. Failing to understand this difference,



the white teacher frequently interrupted the students before they came to the point at which the various elements would tie together. The researchers found that

For the white children in the class, who already have more elements of the schema for topic-centered style, the teacher is better able to collaborate with them and so build on their narrative intentions. With the black children, on the other hand, the teacher's questions lack the rhythmic synchrony and therefore must often be seen by the children as interruptions. Most importantly, the teacher's comments do not build on what the child already knows and so provide the necessary guidance and synchronized collaboration that would lead to the acquisition of an expanded, lexicalized, topic-centered style (p. 658).

Watson-Gegeo and Boggs (1977) found that yet another model for narrative style among native Hawaiian children caused similar evaluation problems.

Cooley's, Michaels and Cook-Gumpers', and Watson-Gegeo and Boggs' studies of oral style all illustrate the importance of ethnographic research for correctly analyzing student behavior upon which assessment is based. In cross-cultural classrooms effective teaching may be blocked by lack of shared assumptions about behavior, leading to student resistance and student failure. Saville-Troike (1980) has characterized the importance of this work: "When the differences are understood, they may be used as an educational base; when they are not, they create a formidable barrier to learning [and assessment, we might add]."

Single culture Studies. In addition to these crosscultural studies, a number of ethnographers have undertaken research in single culture classrooms. They seek to ask the question of how teachers communicate with their students and how they make—or fail to make—their assessment criteria clear. In ethnographic terms, what is the "code" of the classroom society, how do members of this classroom culture acquire the code, and what are the implications of failure to completely command it? This work suggests strongly that teacher evaluation of students may be based not only on the students' comprehension of instructional content, but also on their ability to comprehend behavioral signals.

Gearing and Epstein (1982) undertook a microethnography of a four-student remedial reading group, finding that the teacher's best intentions at fairhandedness and positive motivation for the four students in fact served to further differentiate the two high achieving students from the two low achievers. In the case of this reading group, the key to success for students was the ability to respond not only correctly, but in a precisely timed fashion, judging when the teacher shifted from a monolog mode to a question-asking mode. Given the opportunity to respond to a synonyms game, the two higher achievers were successful, building, as the teacher had hoped, their self-confidence and motivation to learn the words. Looking through the equalizing "chance" factors the teacher had built into the game to make it

"fair" for everyone, the ethnographers were able to observe consistent patterns of minimal or no response on the part of the low achievers. In fact, the game functioned to reinforce their failure. Gearing and Epstein refer to this covert function of differentiation among the students as the "hidden curriculum" of the classroom—a key component of the classroom assessment environment.

Mehan (1979, 1980, 1982) inquired extensively into the structure of classroom events, particularly the shifts in behavior that are required as instructional activities change. Although cast in different terms, he, too, is describing a "hidden curriculum". Success in interactional routines such as turn-taking, he argues, is a critical factor in assessment, regardless of students' mastery of the material. Mehan (1982) contrasts the "content" of classroom instruction with the "form". A student may command the content required by the teacher, but failure to wait one's turn or failure to bid to answer are both assessed negatively.

If a student provides correct content without proper form, the student will be sanctioned. A history of such inappropriate behavior can lead the teacher to treat the student negatively. If a student attends to form without an equivalent concern for content, that student loses opportunities to express knowledge. A history of lost opportunities can lead a teacher to believe that a student is inattentive, unexpressive, and the like. It is in this arena that teachers' expectations are built up, and worked out interactionally. (p. 80)

Nonverbal behaviors have also been the object of ethnographic study. Schultz and Florio (1979) analyzed the social meaning of teacher movement within a monocultural classroom, finding that it functioned as a signal that activity and therefore behavioral expectations for the students were undergoing change. Using videotapes, the researchers correlated a kindergarten teacher's movement within the physical space of the classroom with the changes in her agenda for student activities. In some of the classroom activities, the students were expected to pay close attention to the teacher; in others they were free to continue their private activities. They found that the teacher used physical movement to indicate shift from, for example, announcements to worktime to cleanup time, and that students were evaluated on their responsiveness to these shifts. The researchers found that:

A kindergartner's failure to interpret appropriately the social meaning inherent in the teacher's calls, movement, and use of space can quickly contribute to the formation of a less than promising "institutional biography" for that child. Insight into patterns of interaction operant in the classroom and the children's behavior within them may enable teachers to reflect in a more rich and differentiated way about the children whose performance they are expected to assess (Shultz & Florio, 1979 p. 29-30).

## TOWARD A SYNTHESIS

We claimed at the outset that the field of educational measurement has given too little attention to understanding the assessment issues that teachers face in classrooms on a day to day basis. To illustrate this point, we have reviewed available research on testing to glean insights into classroom assessment. The results of that review were disappointing. To supplement, we turned to emerging research on teacher decision making and available classroom ethnographies for further insights. Still, the composite picture is narrow in scope and out of focus.

### Summary of the Review

Research on Testing. From the body of research on testing in the schools, we know a great deal about large-scale standardized testing for accountability purposes. But in the classroom, we know only that such tests are of secondary importance. We know that nearly all assessments used in the classroom originate with the teacher, and that observation and judgment is an important tool in that context.

Regarding the nature of the assessment process, research studies point out that it is carried out by teachers who typically have little formal training in assessment and what training they have had is narrow in focus. Some data suggest that assessment methods tend to remain the same for a teacher regardless of the purpose, yet assessment methods change as grade and subject matter changes. Research on classroom assessment suggests that both scholastic and social criteria serve in evaluating students, and teachers use such information to form initial and lasting impressions of students. Both students and teachers appear sensitive to those impressions and behave accordingly. One analysis indicates that teachers' paper and pencil tests are short, objective, and measure recall. Another survey suggests that structured performance assessments are used as formal assessments in schools.

However, we know little about the quality of these assessments. One analysis suggests that teacher-developed test items are of poor technical quality. Other studies indicate that teachers do not check the technical quality of their tests and often do not attend to quality control procedures in performance assessment.

We are able to draw only preliminary conclusions about attitudes regarding teacher-developed tests. One study found a few elementary teachers valuing social and family background characteristics more than ability in decision making. Another found teachers generally comfortable with the assessments they use, but concerned about improving their quality. A large-scale survey of students revealed that they find teacher-developed tests more difficult and take them more seriously than standardized tests. In another study, the students reported that (a) tests do not help them know what to study, (b) tests generally call for memorization and (c) they prefer tests requiring short rather than extended answers.



This is not a highly focused picture of teachers' assessment practices. Decades of research on testing in the schools guided by the dominant paradigm tells us little about the classroom assessment environment. Research on teaching and learning—particularly research on teacher decision making—adds some detail to the picture, but not a great deal.

Research on Teaching. First and foremost, this research gives us a first-hand look at the tremendous complexity of the classroom assessment task. Teachers measure dozens of student variables for a variety of reasons and do so at an incredible pace. While planning instruction, teachers tend to focus on activities to be accomplished rather than goals or outcomes. They do consider student characteristics early in the school year while sizing up the class. Those first impressions appear quite stable. During instruction, teachers make interactive decisions about every two minutes. In most cases, they consider student characteristics as antecedents to those decisions. Decisions are also made after instruction (eg., grading, retention, etc.). However, the research on teaching is surprisingly silent on assessment issues in this domain.

Teachers use a variety of strategies to simplify the information processing load of the classroom. Some gather information quickly, form lasting impressions and move on to other tasks. Others focus on a select few, in measuring key student factors and generalize to the class, and still others gain efficiency in measurement by attending to easily counted factors—tasks completed, behaviors exhibited, etc. When measuring achievement, many teachers rely on short, easy to score objective tests of recall. There is little evidence of the impact of these strategies on the quality of resulting data.

Ethnographies of classrooms add a few other details to our picture of classroom assessment. They suggest that discrepancies between the interaction patterns students derive from home cultures and the cultural experience of the teachers can bias assessment results. These studies also suggest that teachers' performance criteria may not always be obvious to students. When students are unaware of key social, scholastic and behavioral criteria, their performance may suffer.

### The Unmet Challenge

Although we know too little about classroom assessment, one point becomes quite obvious: Teachers placed in the environment described above trained only in the measurement methods described in standard measurement texts face real difficulties. There are fundamental and far reaching differences between the science of educational measurement and the assessment demands of the teacher. We have been aware of these differences for decades and generations of researchers have failed to address them. In 1938, Scates was encouraged to see progressive trends moving beyond standardized tests to deeper measurement issues:

The discussion of (contemporary) movements practically fills the testing literature, crowding out much of the former concern over such questions as old-type vs. new-type tests, true/false vs.



multiple choice tests, guessing vs. not guessing... The present concern is with more fundamental issues; the novelty of the objective and standardized tests has passed, the surface attractiveness of carefully printed instruments has worn off, and those workers who are now leading thinking are searching the test movement to ascertain what fundamental values there are (p. 523).

Those educators were unable to reorient thinking. The measurement paradigm in place today was to assume dominance with the accomplishments of psychometricians during World War II.

Five years later, in 1943, Scates clearly articulated some of those fundamental values by pointing out some of the key discrepancies between the science of measurement and classroom assessment. We quote extensively from that paper below, because these passages convey so vividly issues that seem to have been neglected for so long. Scates reveals why we fall short if we train teachers only in the methods of scientific measurement:

It is important to recognize that the criteria of statistical measurement are those of mathematics and the Laboratory; the criteria of the teacher are those which serve to produce good citizens. The two sets differ greatly in emphasis (p. 3).

In the main, science is concerned with abstracting a specific element out of a complex—with isolating a character that is common to a group of objects, and freeing the character from restrictions of immediate circumstance. The teacher's concern is just the opposite. He is working with variable individuals to build a variable product (p. 3).

The scientist may be satisfied with a series of cross-sectional observations; the teacher must be aware of continuing behavior. The scientist is primarily analytical, seeking the elemental, the universal, the permanent; the teacher is primarily constructive, seeking to produce an artistic whole that is unique and changeful (p. 4).

The scientist is seeking truth: Generalization stripped of all complicating factors, which cannot be overthrown by any amount of subsequent investigation... The teacher is not charged with producing a uniform product in toto or in any particular trait... Education is expected to be effective, but not to produce any one thing (p. 5).

The scientist must be strictly uniform, insofar as he can, in his observations. He must have an observational instrument which will reflect the same trait or quality in every instance and in the hands of every observer... The teacher, on the other hand, with much greater tolerance granted him, has no such interest in either objectivity or precision... Impersonal observations may have a more universal quality but they are also more barren... (Thus) the things that science wants out of its observations the teacher wants in (p. 5-6, emphasis in original).

Formal testing cannot be continuous, but the need for watchful observation and interpretation (on the part of the teacher) is continuous... Some tendencies (among students) are revealed by only fleeting manifestations... The teacher must detect changes in attitude, in effort, in cooperative spirit as quickly as they begin to manifest themselves... An author cannot wait until his book sells to write the second chapter, an artist cannot wait until his picture is judged before deciding its composition; and a teacher cannot wait until his pupil is tested before deciding what to do the next hour or the next day (p. 7-8).

To the scientist, the trait measured by a test is uniform throughout its range; to the teacher, growth presents stages of development. This is a serious difference, for a test designed to reflect only changes which can appropriately be represented on a linear scale cannot describe with acceptable fidelity those changes which take place through a series of differing forms of manifestation. New factors appear as each new developmental stage is reached and certain old factors drop out or change in importance....[For example] Reading ability is a complex of many elemental abilities; one stage differs greatly from another, and the factors which it is necessary to evaluate in estimating the satisfactoriness of development in one level may not be factors which it is important to assess when appraising another level (p. 9-10).

The test maker who begins his thinking with the teacher's problems and utilizes technical principles where they will aid, without letting the principles warp his product out of conformity to the practical needs, is bridging the gap between the two situations. But the test maker or the textbook writer who approaches the problem of appraisal by asserting his convictions that all units on a scale must be equal, that a test must be objective, highly reliable, and statistically valid... is just not focusing the practical problem (p. 13).

This review of available research on testing and in fact a review of current training in assessment (for teachers and graduate students) reveals that we have still not met this challenge.

#### Implication for Research: Seeking New Insights

Unanswered questions about classroom assessment abound. Some of them are listed below. When we are able to combine answers to these with generalizable information on topics summarized above, we will be in an excellent position to revise training for teachers, and make measurement a process variable in school improvement efforts:

How do concepts of reliability and validity differ when your goal is to make generalizations within an individual rather than across individuals?

Does student performance vary as the context or assignment varies from classroom to home to work for instance? Or from working alone to working as part of a group? If so, what are the implications for measurement methodology?

Is it possible that interactive decisions (requiring rapid assessment) and post instruction decisions are based on independent but parallel classroom measurement systems? Do data from one often cross over into the others? What are the implications, if reliability and validity standards vary across systems?

What are the crucial differences in assessment environments as grade level increases? As subject changes within and across grade levels? How do teachers adapt to these differing requirements?

How (and how well) are social and personality characteristics assessed by teachers who are left to their own devices? How do these variables weigh in the various pre, instructional, and post decisions? What are the effects on students?

Are teachers initial impressions accurate? If not, how are they changed? What are the effects on students?

Do teachers measure more than recall if we consider the full range of assessment methods they use—not just their paper and pencil tests?

Are teachers short objective tests and quizzes reliable? If not, does an accumulation of unreliable data yield a reliable grade?

Teachers favor activities and content to be covered in planning instruction. Does this translate into student evaluation by counting tasks completed? If so, how is this assessment conducted? How is it translated into feedback?

What are the specific strategies teachers use to simplify the information processing requirements of the classroom? What are the implications of each for the reliability and validity of the results?

What is the assessment process like from the students' perspective? Is it fair? Useful? How does it impact learning? Academic self-concept? Personal self-concept? Does this differ as grade increases, or subject varies? By sex, race social context?

To answer these and other important questions about classroom assessment, a wide range of research procedures must be used, including many not traditionally part of measurement research; stimulated recall, thinking aloud, policy capturing, journal keeping, repertory grids, lens modeling, interviews, group discussions, field observations, case studies with narrative descriptions. The methods are available. Unanswered questions abound. Our research task is clear.

### Implications for Training

As this research is completed, the demands of the classroom assessment environment will become more and more clear. In the meantime, however, it is clear that current teacher and administrator measurement training priorities must change.

Administrators who are currently training to serve as accountability agents by reporting standardized test scores to the school board might also be trained to be instructional leaders—to assist teachers with their day to day measurement of student growth. Such training in classroom assessment methods might be part of administrator certification programs.

Teachers might also be provided with relevant, focused inservice training on classroom assessment strategies and useful quality control procedures. At least some of the content of that training is suggested by the research reviewed here. Training priorities include measuring higher order reasoning skills, writing quality paper and pencil test items, integrating assessment and instruction via oral questioning strategies and designing quality performance assessments—based on observation and judgment.

This new training effort might extend beyond the school walls:

Legislators, taxpayers, parents and the public in general might also be trained to understand the limitations of large-scale testing programs—to understand that the mere presence of a testing program does not assure quality education and that more testing will not of itself produce better schools. The public must become aware of (a) the full range of complex student characteristics (not just math or reading skill) that can be influenced by quality education, and (b) the many alternatives available to us for measuring those characteristics equitably. In short, we need to develop a new generation of critical data consumers—a generation that knows the attributes of good assessment. (Stiggins, 1985, pp 10-11)

### Classroom Assessment: A Key Issue

Why is this issue—an issue disregarded for decades—such a high priority now? There are at least two important reasons.

First, educational outcomes are having greater and greater implications for policy makers. As Cole (1984) and Airasian (1984a) have pointed out, this casts our measures of achievement in new roles and makes it even more crucial that they be of highest quality. It happens that the outcome data given most publicity are the results of standardized testing programs because laypersons and policy makers trust them, yet these measures have been criticized for the narrowness of focus and insensitivity (Dunkin and Biddle, 1974 and Shulman, 1985). Brophy and Good (1985) have reviewed the research on teacher behavior and student achievement and again call for expansion of our conceptualization of achievement to include other measures. We cannot effect such a change until we know more about teacher-developed assessments and find ways to include them in the school effectiveness equation.



A second reason to pursue research on classroom assessment as a high priority is that the improvement of teacher developed assessment can serve as a valuable tool in our school improvement efforts. Measurement can be more than an outcome variable in the effectiveness equation. It can also be a process variable if considered from the teacher's classroom perspective. But classroom assessment will not reach its potential unless and until we focus our research effort on those procedures, discover effective and ineffective practices and then translate research into practice.

## REFERENCES

- Airasian, P.W. (1984a). Educational testing and public policy: Growth and implications. A paper presented at a conference in honor of Benjamin S. Bloom. University of Chicago: Chicago, IL.
- Airasian, P.W. (1984b). Classroom assessment and educational improvement. A paper presented at the conference Classroom Assessment: A Key to Educational Excellence. Northwest Regional Educational Laboratory: Portland, OR.
- Airasian, P.W., Kellaghan, T. & Madaus, G.F. (1977). The stability of teachers' perceptions of pupil characteristics. The Irish Journal of Education, 11(1 and 2), 78-84.
- Airasian, P.W., Kellaghan, T., Madaus, G.F. & Pedulla, J. (1977). Proportion and direction of teacher rating changes of pupil progress attributable to standardized test information. Journal of Educational Psychology, 69(6), 702-709.
- American Psychological Association. (1984). Revised standards for educational and psychological test and manuals. Washington, DC: Authors.
- Brophy, J.F. & Good, T.L. (1970). Teachers communication of differential expectations for children's classroom performance: Some behavioral data. Journal of Educational Psychology, 61(5), 365-274.
- Brophy, J.F. & Good, T.L. (1974). Teacher-student relationships. New York, NY: Holt, Rinehart and Winston.
- Brophy, J.F. & Good, T.L. (1985). Teacher behavior and student achievement. A chapter in M.C. Wittrock (Ed.) Handbook of research on teaching, (3rd ed.). New York, NY: MacMillan.
- Burdin, J.L. (1982). Teacher certification. In H.E. Mitzel (Ed.) Encyclopedia of Education Research (5th ed.). New York, NY: The Free Press.
- Burstein, L. (1983). A word about this issue. Journal of Educational Measurement, 20(2), 99-101.
- Calderhead, T. (1983). A psychological approach to research on teachers' classroom decision making. British Educational Research Journal, 7(1), 51-57.
- Calfee, R.C. & Drum, P.A. (1976). How the researcher can help the reading teacher with classroom assessment. Unpublished manuscript, Stanford University, Stanford, CA.
- Carter, K. (1984). Do teachers understand the principles for writing tests? Journal of Teacher Education, 35(6), 57-60.

- Clark, C.M. & Peterson, P.L. (1985). Teachers' thought process. A chapter in M.C. Wittrock (Ed.) Handbook of research on teaching (3rd ed.). New York, NY: MacMillan.
- Coffman, W.E. (1983). Testing in the Schools: A Historical Perspective. A paper presented at the Center for the Study of Evaluation Annual Invitational Conference, University of California at Los Angeles, Los Angeles, CA.
- Cole, N.S. (1984). Testing and the "crisis" in education. Educational Measurement: Issues and Practice, 3(3), 4-8.
- Cooley, R.E. (1979). Spokes in a wheel: A linguistic and rhetorical analysis of Native American public discourse. Proceedings of the Fifth Annual Meeting. Berkeley, CA: Berkeley Linguistics Society, p. 552-558.
- Dahlof, U. & Lundgren, V.P. (1970). Macro and micro approaches combined for curriculum process evaluation: A Swedish field project. Institute of Education, University of Gotenberg.
- DuMont, R.V., Jr. & Wax, M.L. (1969). Cherokee school society and the intercultural classroom. Human Organization, 28(3) 217-227.
- Dunkin, M.J. & Biddle, B.J. (1974). The study of teaching. New York, NY: Holt, Rinehart & Winston.
- Ebel, R.L. (1979). Essentials of Educational Measurement (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Erickson, F. (1977). Some approaches to inquiry in school-community ethnography. Anthropology and Education Quarterly, 8(2).
- Farr, R. & Griffing, M. (1973). Measurement gaps in teacher education. Journal of Research and Development in Education, 7, 19-28.
- Fleming, M. & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In W.E. Hathaway (Ed.) Testing in the schools. New directions for testing and measurement, No. 19. San Francisco, CA: Jossey-Bass.
- Fyans, L.J. (1985). Teachers as test experts: Hidden talent. A paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Gearing, F. & Epstein, P. (1982). "Learning to Wait: An ethnographic probe into the operations of an item of hidden curriculum." In G. Spindler (Ed.), Doing the Ethnography of Schooling: Educational Anthropology in Action (pp. 240-267). New York, NY: Holt, Rinehart and Winston.
- Gil, D. & Freeman, D. (1980). An investigation of the diagnostic and remedial practices of classroom teachers. Research Series No. 78. East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ED 192 247)
- Good, T.L. & Brophy, J.F. (1978). Looking in Classrooms (2nd ed.). New York, NY: Harper and Row.

- Goslin, D.A. (1967). Teachers and Testing. New York, NY: Russell Sage Foundation.
- Gullickson, A.R. (1982). Survey Data Collected in Survey of South Dakota Teacher's Attitudes and Opinions toward Testing. Vermillion, SD: University of South Dakota.
- Gullickson, A.R. (1984a). Matching teacher training with teacher needs in testing. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Gullickson, A.R. (1984b). Teacher perspectives of their instruction use of tests. Journal of Educational Research, 77(4), 224-246.
- Gullickson, A.R. & Ellwein, M.C. (1985). Post hoc analysis of teacher-made tests: The goodness-of-fit between prescription and practice. Educational Measurement: Issues and Practice, 4(1), 15-18.
- Heartel, E., Ferrara, S., Korpi, M. & Prescott, B. (1984). Testing in secondary schools: Student perspectives. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Heath, S.B. (1982). Questioning at home and at school: A comparative study. In G. Spindler (Ed.), Doing the Ethnography of Schooling: Educational Anthropology in Action (pp. 102-131). New York, NY: Holt, Rinehart and Winston.
- Heath, S.B. (1983). Ways with words language, life and work in communities and classrooms. Cambridge: Cambridge University Press.
- Herbert, G.W. (1974). Teachers' ratings of classroom behavior: Factorial structure. British Journal of Educational Psychology, 44, 233-240.
- Herman, J. & Dorr-Bremme, D.W. (1982). Assessing Students: Teacher's Routine Practices and Reasoning. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.
- Jackson, P.W. (1968). Life in classrooms. New York, NY: Holt, Rinehart and Winston.
- Joyce, B. (1979a). Teachers' thoughts while teacher. Research Series No. 38. East Lansing, MI: Michigan State University, Institute for Research on Teaching.
- Joyce, B. (1979b). Teaching styles at South Bay School. Research Series No. 57. East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ED 187 666)
- Kellaghan, T., Madaus, G.F. & Airasian, P.W. (1982). The effects of standardized testing. Boston, MA: Kluwer-Nijhoff.



- Kuhn, T.S. (1970). The structure of scientific revolution (2nd ed.). Chicago, IL: The University of Chicago Press.
- Labov, W. (1970). The study of non-standard English. Urbana, IL: National Council of Teachers of English.
- Labov, W. (1972). Sociolinguistic patterns. Philadelphia, PA: University of Pennsylvania Press.
- Lazar-Morris, C., Polin, L., May, R., & Barry, L. (1980). A review of the literature on test use. Los Angeles, CA: University of California at Los Angeles, Center for the Study of Evaluation. (ED 204 411).
- Lortie, D. (1975). School teacher. Chicago, IL: University of Chicago Press.
- Marland, P.W. (1977). A study of teachers' interactive thoughts. Unpublished doctoral dissertation, University of Alberta.
- Mehan, H. (1979). Learning lessons social organisation in the classroom. Cambridge, MA: Harvard University Press.
- Mehan, H. (1980). The competent student. Anthropology and Education Quarterly. 11(3), 131-152.
- Mehan, H. (1982). The structure of classroom events and their consequences for student performance. In P. Gilmore and A.A. Glatthorn (Eds.). Children in and out of school. Washington, DC: Center for Applied Linguistics.
- Mehrens, W.A. & Lehmann, I.J. (1984). Measurement and Evaluation in Education and Psychology (3rd ed.). New York, NY: Holt, Rinehart and Winston.
- Michaels, S. & Cook-Gumperz, J. (1979). A study of sharing time with first grad students: Discourse narratives in the classroom. Proceedings of the Fifth Annual Meeting. Berkeley, CA: Berkeley Linguistics Society (p. 647-660).
- Morine-Dershimer, G. (1979). Teacher conceptions of children. Research Series No. 59. East Lansing, MI: Michigan State University, Institute for Research on Teaching.
- Morrison, A. & McIntyre D. (1969). Teachers and teaching. Harmondsworth, Middlesex: Penguin Books.
- Pedulla, J.J., Airasian, P.W. & Madaus, G.F. (1980). Do teacher ratings and standardized test results of students yield the same information? American Educational Research Journal, 17, 303-307.
- Peterson, P.L. & Barger, S.A. (in press). Attribution theory and teacher expectancy. In J.B. Dasek (Ed.) Teacher expectancies. Hillsdale, NJ: Laurence Earlbaum.

- Philips, S.U. (1972). Participant structures and communicative competence. Warm Springs children in community and classroom. In C.B. Casden, V.P. John and D. Hymes (Eds.) Functions of Language in the Classroom. New York, NY: Teachers College Press.
- Prawat, R.S. (1979). Teacher perceptions of student affect. Research Series No. 44. East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ED 173 326)
- Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. Harvard Educational Review, 40, 411-451.
- Rudman, H.E., Kelley, J.L., Wanous, D.S., Mehrens, W.A., Clark, C.M. & Porter, A.C. (1980). Integrating assessment with instruction: A review (1922-1980). Research Series No. 75. East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ED 189 136)
- Salmon-Cox, L. (1980). Teachers and tests: What's really happening? Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.
- Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? Phi Delta Kappan, 62, 631-634.
- Saville-Troike, M. (1980). Cross-cultural communication in the classroom. In J. E. Alatis (Ed.) Georgetown University Round Table on Languages and Linguistics, (pp. 348-355). Washington, DC: Georgetown University Press.
- Scates, D.E. (1938). The improvement of classroom testing. Review of Educational Research, 8(5), 523-536.
- Scates, D.E. (1943). Differences between measurement criteria of pure scientists and of classroom teachers. Journal of Educational Research, 37, 1-13.
- Schultz, J. & Florio, S. (1979). Stop and freeze: The negotiation of social and physical space in a kindergarten/first-grade classroom. Occasional Paper No. 26. East Lansing, MI: Michigan State University, Institute for Research on Teaching.
- Shavelson, R.J., Cadwell, J. & Izu, T. (1977). Teachers' sensitivity to the reliability of information in making pedagogical decisions. American Educational Research Journal, 14(2), 83-97.
- Shavelson, R.J. & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions and behavior. Review of Educational Research, 41(4), 455-498.
- Shulman, L.S. (1980). Test design: A view from practice. Chapter 5 in E.L. Baker and E.S. Quellmalz (Eds.) (pp. 63-73) Educational testing and evaluation. Los Angeles, CA: Sage.

- Shulman, L.S. (1985). Research programs for the student of teaching: A contemporary perspective. A chapter in M.C. Wittrock (Ed.) It and back of research on teaching (3rd ed.). New York, NY: MacMillan.
- Simon, H.A. (1957). Models of man: Social and rational: Mathematical essays. New York, NY: Wiley.
- Sproul, L. & Zubrow, D. (1982). Standardized testing from the administrative perspective. Phi Delta Kappan, 628-631.
- Stetz, F. & Beck, M. (1979). Comments from the classroom: teachers' and students' opinions of achievement tests. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Stiggins, R.J. (1985). Classroom assessment: A key to instructional quality. Educational Leadership,
- Stiggins, R.J. & Bridgeford, W.J. (1985). The ecology of classroom assessment. Journal of Educational Measurement, 22(4).
- Stinnett, T.M. (1969). Teacher certification. In R.L. Ebel (Ed.) Encyclopedia of educational research (4th ed.). New York, NY: MacMillan.
- Tollefson, M., Tracy, D.B., Kaiser, J., Chen, J.S. & Kleinsasser, A. (1985). Teachers' attitudes toward tests. A paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Ward, J.G. (1982). An overview of the AFT's teaching and testing. In S.B. Anderson and L.B. Coburn (Eds.) Academic testing and the consumer (pp. 47-52). San Francisco, CA: Jossey-Bass.
- Watson-Gegeo, K.A. & Boggs, S.T. (1977). From Verbal play to talk story: The role of routines in speech events among Hawaiian children. Child Discourse, S. Ervin-Tripp (Ed.), (pp. 67-90). New York, NY: Academic Press.
- Weiner, B. & Kukla, A. (1974). An attributional analysis of achievement motivation. Journal of Personality and Social Psychology, 15, 1-20.
- Weinshank, A.B. (1980). Investigations of the diagnostic reliability of reading specialists, learning disabilities specialists, and classroom teachers: Results and implications. Research Series No. 72. East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ED 189 574)
- Whitmer, S.P. (1983). A descriptive multimethod study of teacher judgment during the marking process. Research Series No. 122. East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ED 234 052)
- Woeller, E.H. (1979). Requirements for certification for elementary school, secondary school and junior college teachers. Chicago, IL: University of Chicago Press.

Yeh, J. (1978). Test use in the schools. Los Angeles, CA: University of California at Los Angeles, Center for the Study of Evaluation.

Yinger, R. (1977). A study of teacher planning: Description and theory development using ethnographic and informal process methods. Unpublished doctoral dissertation, Michigan State University.

Zaharik, J.A. (1975). Teachers' planning models. Educational Leadership, 33, 134-139.



Chapter 3  
Ethnographies of Classroom Assessment

ABSTRACT

Three case studies of classroom assessment environments are summarized in this chapter. Each is described in terms of background information, assessment schedules and procedures and teacher and student perspectives on assessment. All three studies focused on 6th grade classrooms to allow for some comparability of results. Study one describes a classroom in a middle school in a small industrial community. Study two examines a classroom in an affluent suburban community. And study three explores assessment in a private urban school.

## CASE STUDY #1 Background Information

### The Community

The information included in this classroom assessment environment case study was gathered via observations in a 6th grade classroom in a middle school in a small industrial community during April and May of 1985. The population of the community is approximately 5500 residents. It is located about 25 miles from the nearest metropolitan area. The community supports a high school, the middle school and two elementary schools.

### The School

The school staff includes a principal, vice principal, 28 teachers, 2 secretaries, 4 classroom aides, a librarian and 2 counselors. Total enrollment is 533. Of these, 151 are sixth graders divided into 6 classes, one of which is taught by Carol. Carol's classroom was the focus of this case study.

### The Teacher

The 1984-85 academic year marked Carol's return to full-time teaching after an absence of 10 years. She was trained at a major midwestern university, taught for two years in an open community school, taught primary grades for 2 years and then left teaching. After several years as a social worker, Carol returned to education, serving as a substitute teacher for one year before returning to the classroom full time in the fall of 1984 where she was placed in charge of 27 sixth graders for the 1984-85 school year.

### The Classroom Routine

It was Carol's practice to manage the classroom in a routine manner. The daily routine was as follows:

7:45-8:40 Mathematics. The students arrive and begin immediately to work on 2 to 4 "warm up" math problems that Carol placed on the board before their arrival. They solve and discuss the problems and then turn attention to grading homework. Students trade papers, Carol reads answers, students count the number wrong, Carol places the letter grade scale on the board and students grade and return papers. Papers are then handed in. About 15 minutes into the class period, instruction begins on new material. This often involves recitation and student work on the board, after which students are given their assignment and 10 to 20 minutes to work on it. During assignment work time, Carol often identifies students who had difficulty with the previous day's assignment and gathers them in the back of the room for special help.

8:45-9:40 Science (M, W, F) or Computers (T, Th). Science class also begins with the grading of homework. Students trade papers, answers are provided by Carol or class members, grades are assigned and papers are handed in. Occasionally, students read their grades aloud for entry into the grade book. This is followed by about 30 minutes of instruction (experiment, project work, reading from the text, discussion, etc.) and 5 to 15 minutes to work on the homework assignment.

Computer instruction is completely individualized, with each student working at a terminal. Since there are twice as many students as terminals, students work with partners, dividing terminal time. Instruction and evaluation are completed unit by unit on the terminal, where progress records are stored. Carol offers assistance as needed and checks unit completion records before the student can proceed to the next unit.

9:45-10:20 Planning Period. During this time, students participate in a variety of activities outside the classroom. Some go to physical education, some to band, others to work on special projects. Carol uses the time for planning, materials preparation, etc.

10:25-11:10 Language Arts. Students begin this period with a 10 to 15 minute journal writing assignment as specified by Carol. Occasionally, journal entries are shared with the class. Students then trade workbook homework papers and grade them using the standard percent correct scale, return them, and hand them in. Instruction on new material is carried out, an assignment is given and students work on their assignment.

11:10-11:30 Spelling. Students trade homework papers, correct and grade them and hand them in. They then receive an assignment which they begin immediately to work on. One day per week during this time slot is spent in the library.

11:30-12:00 Lunch.

12:10-12:25 Home Room. This is designated as personal, recreational reading time. Carol reads to the class and they read books of their own.

12:30-1:25 Reading. Students are ability-grouped for reading instruction. As soon as those from other classes arrive, the students trade homework assignments, grade them and hand them in. They then work on the next unit of reading instruction or activity, which often includes completion of part of their next reading assignment.

1:25-2:15 Social Studies. Students trade, grade and hand in homework. Instruction follows, typically involving the entire class in discussion and recitation, movies, reading from the textbook or some combination of these. They are then given an assignment which they work on until the day ends at 2:15.

Variations in Routine. While this is a description of a typical day in Carol's class, there were occasional variations. For instance, some students were called out of some classes on a daily basis for special instruction. In addition, special activities, such as movies, speakers, etc., often required that sixth grade classes be combined for a period and on occasion the entire class leaves the room or school for a period of time during the day for field trips, to view exhibits, etc. Also, unit tests and other assessments were common and these broke up the routine described above. These were only occasional variations in what turns out to be the highly structured environment that is central to Carol's classroom management scheme.



## The Assessment Environment

As I watched this routine with its variations unfold through 14 days of observation over a six-week period, I was intent on finding, documenting and understanding those events that comprised the assessment environment in this classroom. I sought to understand the purposes that drive the assessment system from the teacher's and students' point of view. I sought to understand what student characteristics are measured, using what measurement strategies, to create what kinds of records and feedback. And finally, I tried to understand the impact of that environment on student learning and attitudes about learning. In the report that follows, I will collect the events and insights derived from extensive participant observation into a portrait of one particular assessment environment.

### The Purposes for Classroom Assessment

The purpose that dominates classroom assessment in this case is the assignment of grades. The specific procedures used to achieve this purpose will be described later. But clearly, a vast majority of the assessment practices feed into the grading decision. Nearly every sample of written student work is transformed into a grade. That means 8 to 10 work samples (e.g., assignments, tests, etc.) across all subjects per day nearly every day are assigned grades and those grades are recorded for later averaging.

However, there is also clear evidence that grading is not the only purpose for classroom assessment. For instance, Carol uses homework to diagnose which students had difficulty with the assignment. In most class periods immediately after homework or tests are handed in, while students work on the next day's assignment, she scans the grades, identifies those who had problems and gathers them in a backroom for special help. This is a daily routine.

I also observed some diagnosis of group needs, but this is not as frequent. On occasion, such as when everyone did poorly due to inadequate instruction from a substitute, Carol recycled, covering and regrading the same material. But by and large, the need to keep up a constant pace of instruction in order to cover the material required for the year kept Carol from being able to recycle the entire group too often. Rather, she chose to recycle individuals. Another illustration of this was seen in reading. Students who failed to reach mastery on specific parts of criterion referenced unit tests were given individual instruction on those parts and retested later. The Beacons reading series (Houghton Mifflin) was structured to allow this. Further, Carol always computes average grades for students at mid-trimester, even though they are not required or reported, because those average grades gave her a sense of the individual needs and progress of students.

Carol also reported using her own assessments early in the year to size students up—to estimate achievement potential. She did not consult prior student records for this purpose, but chose instead to rely on her own data and insights. She reported being able to make relatively accurate judgments

of potential based on 3 to 4 weeks of assignments, tests, and behavioral evidence of self-management and instruction. This year's first impressions were borne out over the year, with the exception of one student, whom she misevaluated. This was a new student in the community who made an outstanding first impression which was not sustained.

Another obvious use of classroom assessment in this environment is to control and motivate purposeful student behavior. This control takes two forms, one subtle and one obvious. Subtle control comes from the use of assessment to communicate achievement expectations. This was done through example and through oral questions and answers during class. For example, Carol used examples of good and poor work to reinforce the level of performance expected. She also communicated the levels of cognitive operation she expected with her question and answer strategies during instruction. The extreme importance of these levels in the environment is addressed later. But clearly, students learned their achievement targets from Carol's assessments; not from lists of instructional objectives she handed out.

The more obvious control comes from the heavy emphasis on grading and completing work in order to receive high grades. As mentioned above, nearly all student work is graded. A chart of incomplete assignments is posted for students to review. The opportunity to participate in a special field trip is the reward for work completion. When grades are assigned, students always have the option of not reading a low grade into the book. They can redo the work and "go for a higher grade." Without question, students are motivated through grade practices.

Finally, assessment serves the purpose of providing feedback to parents beyond grades assigned. Based on classroom assessments, teachers are able to select from among a predetermined computerized set of 100 alternative comments they want to communicate to parents about a student's work. These comments are then printed on the report card by the computer along with the grade.

Another use of assessment uncovered during my visits was as an instructional strategy. Both science and social studies textbooks (and accompanying teacher guides) used questioning during instruction to trigger student thinking and learning. For instance, when a key point was made or a key concept defined in the text, students were immediately asked a series of questions to help them internalize the key point. In addition, Carol occasionally gave tests under apparently real test conditions and then didn't take grades, preferring rather to tell the students that the event was intended as a learning experience. These are examples of the tight coupling of assessment and instruction to promote student learning.

During my observations, I was also able to find evidence of the systematic use of assessment for ability grouping in math and reading, selection for advanced and remedial programs (involving very few students) and providing information on student achievement to school managers via the standardized achievement test battery. However, none of these purposes was served by Carol's classroom assessments. All were based on testing programs carried out outside the classroom context using published tests. There is a district wide standardized testing program. The achievement battery was administered prior to my observations. Carol received the results during the case study. She showed interest in them and put them to no apparent use.

In sum, the classroom assessment environment served many purposes. One purpose—grading—was clearly dominant. But many other important purposes were also given attention. Most of those relate to the promotion of individual student learning.

### Student Characteristics Assessed

The classroom assessment environment in this case focused on a wide variety of vastly different student characteristics. Of course, the primary and most visible focus was student achievement. But even the measurement of this most important outcome was carried out with surprising dimensionality. I will illustrate that point below. But in addition, other student characteristics assessed included student ability, classroom behavior and social development and personality.

Achievement. Turning first to the heart of the assessment matter, based on my observations, it was clear that students were assessed to determine mastery of a wide range of content knowledge and a wide range of reasoning skills. Let me begin by exploring the assessment of reasoning skills. Table 1 reports the results of an analysis of samples of questions posed to students in different sources within each subject. Samples of question from each subject matter area were classified according to the level each represented on Bloom's Taxonomy of the Cognitive Domain. Note that all subjects probed far more than the recall of facts and information. Note also the high degree of consistency in percentages across different sources within subjects—with the important exceptions of the total mismatch in social studies and slight misalignment in science. This consistency suggests that assessments reflected what was taught, at least with respect to reasoning skills in math, reading and science. The social studies assessments lacked instructional validity. Carol had been vaguely dissatisfied with the tests before seeing the results of this confirming analysis.

However, among the levels of reasoning skills tapped, one was noticeably absent. Rarely were students asked to make and defend a value judgment—to evaluate. Such questions are rarely posed in texts, by Carol in class (she posed evaluation exercises twice during the 14 days of observation) or on assessments. The one exception to this was a series of evaluative questions posed on a criterion referenced reading test—on which students scored uncharacteristically low. Further, when Carol asked for evaluative comments on a movie shown in class, students were reticent to express their opinions. Evaluative reasoning is given little attention in this classroom.

**Table 1**  
**Analysis of Depth of Questioning**

**Social Studies**

|               | <u>Workbook</u> | <u>Text</u> | <u>Test Review Sheet</u> | <u>Assessment Unit Test</u> |
|---------------|-----------------|-------------|--------------------------|-----------------------------|
| Recall        | 21%             | 16%         | 79%                      | 84%                         |
| Comprehension | 30%             | 41%         | 3%                       | 0                           |
| Application   | 17%             | 0           | 0                        | 0                           |
| Analysis      | 10%             | 19%         | 18%                      | 16%                         |
| Synthesis     | 10%             | 24%         | 0                        | 0                           |
| Evaluation    | 13%             | 0           | 0                        | 0                           |

**Science**

|               | <u>Activity Book</u> | <u>Text</u> | <u>Worksheet</u> | <u>Unit Test</u> |
|---------------|----------------------|-------------|------------------|------------------|
| Recall        | 4%                   | 26%         | 8%               | 50%              |
| Comprehension | 35%                  | 26%         | 24%              | 15%              |
| Application   | 21%                  | 7%          | 32%              | 10%              |
| Analysis      | 18%                  | 26%         | 24%              | 25%              |
| Synthesis     | 14%                  | 11%         | 4%               | 0                |
| Evaluation    | 7%                   | 4%          | 8%               | 0                |

**Math**

|               | <u>Text</u> | <u>Class</u> | <u>Assessment</u> |
|---------------|-------------|--------------|-------------------|
| Recall        | 0%          | 0%           | 0%                |
| Comprehension | 0%          | 5%           | 0%                |
| Application   | 88%         | 86%          | 81%               |
| Analysis      | 12%         | 9%           | 19%               |
| Synthesis     | 0%          | 0%           | 0%                |
| Evaluation    | 0%          | 0%           | 0%                |

**Reading**

|               | <u>Text</u> | <u>Teacher Manual*</u> | <u>Practice Book</u> | <u>Test Manual</u> |
|---------------|-------------|------------------------|----------------------|--------------------|
| Recall        | 0%          | L 55%                  | 0%                   | 4%                 |
| Comprehension | 67%         |                        | 41%                  | 50%                |
| Application   | 19%         |                        | 11%                  | 8%                 |
| Analysis      | 14%         | I 24%                  | 30%                  | 29%                |
| Synthesis     | 0%          |                        | 0%                   | 8%                 |
| Evaluation    | 0%          | E 21%                  | 18%                  | 0%                 |

\*Discussion questions grouped as L=understand or remember, I=interpret, E=evaluation.



Nearly all assessments are like the reading test in that they are interpreted in a criterion-referenced manner. In fact, I can recall only one instance during my observations when an assessment served a norm-referenced purpose—that is, to rank students. That was for placement purposes in the ability grouping context. In virtually every other case, fixed cutoff scores were established for the assignment of grades, so theoretically, every student could receive an A on any given assessment.

Without question, these assessments measured how much students knew and how well they were able to use their knowledge. Proceeding through the daily schedule, in math students were called upon to apply what they had learned to solve problems; in science, all levels of reasoning skills were required; in computers, students were called upon to perform and show progress in computer operations. In language arts, students performed in two distinct ways, but only one was evaluated. They wrote daily in their journals, but their writing was almost never formally evaluated. During my observations, day-to-day instruction focused on learning 16 rules of capitalization. Application of these rules was consistently evaluated and graded. Spelling was dominated by assessment of ability to recall correct spelling. Reading assessment focused primarily on mastery of comprehension objectives, with application and analysis also given attention. And social studies assessments varied greatly. While Carol used recitation to assess the full range of reasoning skills, the tests provided with text materials tested only recall of facts and information, as previously mentioned.

In carrying out the paper and pencil assessments (e.g., unit test, quizzes, etc.) Carol relied completely on assessments provided by the publishers of the instructional materials. She developed none of her own assessments. Since this was her first year with new materials in sixth grade and since she faced 6 new preparations daily, time did not permit new test development.

Another dimension of student achievement that deserves prominent mention in this discussion is the assessment of students' pace of work completion. Teachers at this school are very concerned about students completing work on time. Students loose one grade level for every day work is handed in late. This has had the effect of making some students supremely aware of the need to work at the fastest possible pace—regardless of the quality of the result. The students first instinct when they began computer work was to move ahead as rapidly as possible, until Carol slowed them with admonitions that they attend to quality. When students do board work in math, they compete to be the first to solve the problem, in many cases regardless of accuracy. In letters to imaginary fifth graders advising them how to get good grades in sixth grade, the most frequent suggestion among Carol's students was to get work in on time. The same was true of the students in a neighboring classroom. Students urged their younger friends to try hard and listen and behave in class. But doing high quality work was rarely mentioned.

In sum, the measurement of achievement in this case includes both an assessment of how much students learn and whether they complete their work in a timely manner. To an outside observer, the latter sometimes appears to outweigh the former.

Ability. Carol gathers and uses information on student ability. This assessment is subjective, multidimensional and very important. It is subjective in that ability is defined as the amount of effort expended by the student and the pride students take in their work in relation to the actual quality of achieved results. Carol observes students across a variety of subjects over an extended period of time. She uses personal interactions and assignments to track how long it takes them to master new concepts and how much difficulty they have in doing so. Some understand immediately, according to oral and written classroom assessment, while others may understand by the end of a period of instruction and still others may take several days to understand. In a matter of weeks and months, she sees the pattern emerge for each individual student.

This is an important assessment because it allows her to establish expectations for each student. The student is evaluated in part in terms of perceived ability and achievement in relation to expectations. But in this case, I don't mean evaluation in terms of grades. Those are almost totally determined by percentages correct on and timely completion of tests and assignments. But if a student's average is borderline, work in relation to expectations is considered. These expectations are also considered by Carol in choosing comments to enter on the report card and in conducting day-to-day classroom interactions. Carol may not be specifically aware of it, but I think she poses questions to students that will allow them to feel some sense of success—very tough questions occasionally to the one very bright girl in the class (that only she could answer) and less challenging questions to less able students, whom Carol calls on to answer from among many volunteers. This aspect of the classroom assessment environment was subtle in its manifestation and very sensitive in its treatment of students.

Classroom Behavior and Social Development. While ability, achievement and pace of work are central aspects of Carol's assessment environment, other student factors come into play also. For instance, classroom behavior and compliance with specified rules are carefully monitored. Expectations are made clear both in terms of verbal descriptions of appropriate behavior and in terms of feedback for inappropriate behavior. That feedback is delivered swiftly in a very private, personal manner—never involving public embarrassment for the perpetrator.

Carol reports that she monitors social development in terms of peer relations, group work, temperament, activities during non-academic times and her interactions with them. She also draws inferences about personality in these same ways, adding confidence, anxiety, self-concept and sense of humor to the traits observed. These student characteristics are not trivial. They are central to effective classroom functioning—as crucial as any grade or test score. This fact was made abundantly clear during my observations. Let me illustrate.

This classroom environment is managed via a relatively tight routine; rules of behavior and expectations are clear. Violations of accepted personal and social interaction patterns—which are defined by the teacher—are also enforced by the teacher. The importance of the evaluation and sanctions that hold this system together became very clear with the arrival of a substitute

teacher. The tightly managed interpersonal routine was broken and the environment immediately deteriorated to near chaos. However, when Carol returned, she was able to reassert the routine, her expectations and sanctions to restore order. It became apparent that the quality of a classroom learning environment hinges to a great extent on evaluation of student social behavior.

### Assessment Methods

Student achievement was assessed in many different ways in this classroom. These included paper and pencil assessments (including homework assignments), oral questions and observations of student behavior and products (i.e., performance assessments). As mentioned previously, virtually all of the assessments in the first two categories were based on exercises provided by the publishers of the instructional materials used in Carol's classroom. However, oral questions and performance assessments tended to be products of Carol's planning.

Assessments by Subject. These assessment methods varied only a bit from subject to subject. Student achievement in science was measured via performance on daily text and activity book assignments and unit tests provided by the publisher. The daily assignments required the student to answer a series of questions over material covered in text chapters, or they included workbook exercises related to experiments conducted during class time. Unit tests included about 20 multiple-choice test items along with a few fillin items.

The oral questions Carol posed during class arose from several sources, depending on the mode of instruction that day. If an experiment was conducted, the teacher's manual provided probes for use during class. If instruction revolved around reading the text, it included discussion questions as part of the presentation. In addition, Carol generated most of the follow up questions that direct class discussion. I will say much more about her particular questioning strategies in a separate section below.

Performance assessments also played a role in science assessment, as students were occasionally called upon to produce certain products, such as to use common household materials to construct a model of an atom. Performance assessment of a different variety provided the sole basis of assessment in computer science. In this case, students had to carryout the steps in the computer operation process to the computer's satisfaction. These assessments focused entirely on keyboarding skills.

Math assessment was the most unidimensional. Assessments were based on student ability to solve computational and story problems. Board work, oral questions, textbook assignments and chapter tests all focused on these same skills.

Language arts assessments were based on homework workbook assignments that required proper application of the grammar, punctuation and capitalization roles that seem to dominate instruction at this grade level. The tests observed focused on memorization of 16 rules for using capitalization. These



assessments were brief and sharply focused on specific skills, as were oral questions during class. Performance assessments of two sorts were observed. Students were called upon to write. But evaluation of their writing was informal--almost casual. When it was formal, the formality focused on the correctness of the presentation more than the quality of the prose. A second kind of performance observed was the actual construction (writing, illustrating and binding) of a book by students.

Spelling assessment, like math, was straight forward. Spelling worksheets and tests tested the ability to spell correctly and find correct misspelled words in lists.

Reading assessments included many more dimensions. Text and workbook assignments called for the application of a variety of different reading skills. Most assignment exercises and unit test items were multiple-choice or fillin. Unit tests were criterion (specific objective) referenced, including five items per objective and up to 10 objectives per test. Mastery level on these power tests was 3 of 5 items correct. Objectives not mastered were retested at a later time, using parallel tests. Many of the oral questions that guided instruction were open-ended and were provided by the teacher's manual. Performance assessments took the form of required written book reports, due every 6 weeks. Performance criteria, scoring standards and alternative strategies for satisfying this requirement were spelled out in writing for students.

Written assessments in social studies were based on multiple-choice and fillin items on homework assignments and unit tests. The oral questions which direct class instruction were posed in the text presentation. Homework assignments included workbook sheets and responses to items in the text.

Similarities Across Subjects. There are some striking similarities across these different subjects. First, everything was graded. More about these procedures below. But clearly, grades dominated the assessment environment by brute force and sheer numbers alone. Second, almost all assessments (oral and written) were based on exercises for which the response is either correct or incorrect. Shades of correctness almost never came into play. This pattern relented somewhat in performance assessments, where degrees of quality come into play. Somehow the rigidity of this correct/incorrect dimension of the environment may relate to the students' difficulty in making their own evaluative judgments. The rewards of the environment came to students who knew the answer. Matters of opinion received far less attention.

A third important similarity across subjects not mentioned previously, was that students were frequently called upon to do self-assessments. They had to take responsibility for letting Carol know when they needed more or special help. Early in the year, she reported that they were less willing to risk exposing their weaknesses. But after discussing the advantages and seeing the risk pay off many had become open to the idea. For example, when special help was offered in the backroom to those receiving low grades on assignments from the night before, some students who had received good grades were always present.



Grades and Grading Practices. Carol made her expectations most clear through her grading procedures. Those procedures were fixed across all subjects. Ninety percent correct received an A, 80 percent a B, 70 percent a C, 60 percent a D and below 60 percent an F. This sequence applied to the longest or shortest assignment or test. Even in those performance assessment cases where percent correct does not apply, the continuum is covered to grades for recording. Grades were entered in the book daily. All entries were given equal weight (regardless of length or importance) and the accumulated grades were averaged at the end of the trimester.

The domination of grading practices can only be seen when considered in light of the instructional time spent. Homework assignments were graded a minimum of 6 times per day (once per period) at the start of each period. Assuming that it takes 10 minutes per period 6 periods everyday to trade papers, grade them and read scores, that means an hour per day was taken up grading homework. When we add tests, quizzes and performance assessments, we can rapidly accumulate a quarter of the day being given to activities directly related to grades.

Oral Questions. These tend not be related to grades, but did contribute to the nature of the classroom assessment environment. Carol asked many questions during the day. I estimate that questions were asked at a rate of 30 to 40 per period over 6 periods. That means that the attending child needs to be ready to answer 180-240 questions per day. Obviously, no child attends continuously. But some definitely attend more than others. To find out who, I observed and charted Carol's question asking patterns.

The first observations attempted to track Carol's "action zone" in terms of room arrangement. The original research on this topic found teachers interacting most with students in the center of the room and less with students in the extremes. Carol's pattern differs. Her average number of questions per student per period was twice as high on left and right sides of the room as down the middle. This might allow kids in the middle to rest somewhat more.

The second set of observations focused on who Carol called on to respond. She called on volunteers over 90 percent of the time and non-volunteers the rest. Volunteers answered correctly at least 80 percent of the time, while non-volunteers were right less than 30 percent. Apparently volunteers were tuned in, while non-volunteers had learned that they could rest if their hand was not in the air. Carol revised this procedure upon learning these results and the percent of correct responses by non-volunteers began to increase.

Performance Assessments. While paper and pencil assessments and oral questions clearly dominated the assessment environment, student product and behavior observations and ratings also played a key role. Examples of performance assessments at work were mentioned above. These included writing assessments, book reports, science product assignments and computer keyboarding skill assessments. Others included production of social studies maps, art products, evaluations of student classroom social behavior and performance tests in music and physical education.

What was most striking about these assessments was the range of clarity with which performance expectations or standards were communicated to students. In some cases the requirements were crystal clear. The computers were always clear in what they expected both in pace and quality of performance. Carol was similarly clear in communicating expectations about book reports and classroom social behavior. However, she was consistently less clear in stating what she meant by good journal writing, good science products, quality social studies maps, etc. And in these cases, when grades came into play, she tended to rely on easily countable (correct/incorrect) features of the product, such as mechanical or appearance problems in writing or accuracy of labels on maps. When reversion to the objective criteria made no sense, she tended not to assign grades to the assessments. In short, I would estimate that students were very clear on Carol's performance expectations in about half of the performance assessments observed.

Summary of Assessment Methods. The important generalizations arising from these observations of assessment methods are these: In this environment, paper and pencil assessments and oral questions dominate, but service difference purposes. Paper and pencil assessments—nearly all of which include items that are correct or incorrect—feed into grade procedures. These items are prepared by the text publishers. Oral questions, on the other hand, rarely feed into grades, but do serve to motivate students to attend and learn. Another kind of assessment—performance assessment—fills the middle ground between the two. Students seem more motivated to involve themselves in performance-based measurement activities than paper and pencil tests, even though these too feed into their grades.

#### Feedback Arising from Assessment

Throughout the foregoing discussion, I have commented on various aspects of the feedback given to students and parents on student development. Some examples included the use of grades and computer printed comments on report cards and the swift but very private delivery of feedback by Carol to students whose classroom social behavior is inappropriate. However, I was able to make some additional observations about the feedback dimension of this assessment environment that deserve comment.

For example, I found marked differences among students in their need for feedback. In computer class, some students repeatedly volunteered information on their achievement, seeking praise from Carol. In other classes, some students repeated and continuously volunteered to respond to questions and do board work, while others tended not to risk. Whether this is due to differences in knowledge level, confidence attention or some combination of these is not apparent. But this is a central aspect of the assessment environment about which we need to know more. It is obviously a reward to be called on to answer, because it means to the student Carol has confidence in her or him to respond correctly. The motivational potential of this factor would be immense if it could be harnessed.

I also want to comment further on how Carol delivers feedback to students and parents. Grades, of course, represent the primary mode of communication. But there are others. She used the comments on report cards, but found the computerized list of 100 possibilities inadequate and impersonal, particularly for the average child. She used group praise extensively, commenting on how well the class did as a group on certain tests and assignments. She also consistently praised correct answers to questions during recitation. These represented key aspects of this classroom assessment environment.

I have already commented on the kinds of performance that are rewarded but want to reinforce one key point in that regard. The single most prominent piece of feedback observed during my term of observation was a field trip to a local restaurant for lunch and video games for those who had completed all of their assignments. This had the effect of greatly reinforcing the desirability of completing work. I may have overlooked other important aspects of feedback, but I saw no such visible feedback for doing quality work. The assumption may be that grades are the reward for that aspect of performance. If so, it might be interesting to compare the perceived value of the two forms of feedback in the students' eyes.

#### Student Views of the Assessment Environment

In order to gain some understanding of how the students view the assessment environment in Carol's classroom, I talked with them casually, listened to them interact with one another and interviewed several of them. As a result of these observations, I am able to draw the following conclusions:

- Students see grades and grading practices as fair. They know what is expected and they know that those who don't study don't get good grades.
- Students know they can always go for a better grade on an assignment but rarely do so. Time rarely permits them to redo the assignment immediately. And, after a day or two, they lose the motivation to redo the work.
- Good students see themselves in control of the rewards they receive, while less able students tend to attribute responsibility to others (teachers or administrators). They tend not to take responsibility for their poor performance.
- Carol often has students read grades aloud to enter in the grade book. She does this for efficiency. The effect is that it makes grades all that much more dominant in the environment. Students don't like having to read low grades.
- Some slower students remain reticent to let Carol know their weaknesses for two reasons--they don't like to be the one slowing everyone else down, and/or they don't want others to know.

- o Students vary in how they prepare for big tests. Most rely heavily on workbook and text review sheets. Many strive to memorize material and have parents ask them trial questions. Good students have better preparation strategies than poor students. Students see tests as problems in memory and little else.
- o Students are very aware of the constant barrage of grades they receive day after day. They often find it tiring and sometimes difficult to gain control over. The constant demanding pace makes it difficult to keep quality up--particularly for poorer students.

There is no question that student perceptions of the assessment environment vary little with the students position on the achievement continuum. Good students have few difficulties and would make few changes in the assessment processes. Poor students often have unusual perceptions of what it takes to be successful in the environment. For instance, one student had concluded, after studying hard and failing and not studying and doing well (once or twice), that it was pointless to prepare for tests. A key to helping these students to achieve better may be to try to learn and correct their inappropriate, ineffective study strategies. In any event, there may be specific aspects of the classroom assessment environment that present unique problems to the less able student. The effect of these may be to perpetuate some of the poor performance. We need to study and understand these more.



## CASE STUDY #2

### Background Information

This report describes classroom assessment practices in a suburban sixth grade classroom. Information for the study was acquired via classroom observation and teacher interviews during the spring of 1985. The study's primary purpose is to provide initial and specific classroom information on a range of practical assessment issues, including how teachers use assessments, what techniques they rely on, what student characteristics they evaluate, how they record assessment results and provide feedback to students and others. In the following pages, I first provide a brief description of the school and general classroom setting and schedule and then describe, in greater detail, the assessment practices in that setting.

#### The School and Classroom Environment

This sixth grade class was located in an attractive, suburban elementary school in Oregon. The school, a modern single-level building, was situated on a quiet street in the midst of a rapidly growing and predominantly professional community. The largest of three elementary schools in the community, Adam elementary school included about 500 students from kindergarten through sixth grade. The classroom I observed was one of two sixth grade classes, each with approximately 26 students. Although each sixth grade teacher instructed a core group of students, neither classroom was fully self-contained. Students were grouped by ability for both math and reading/spelling. Ann Larson, the teacher in this classroom, was responsible for instructing her homeroom students in English and social studies as well as teaching a combined group of 6th graders in reading/spelling and math.

Students in the class were evenly divided between boys and girls as well as between low and high performers. In fact, students in Ann's homeroom class ranged substantially in ability. Eleven of the students, close to half of the class, were assigned to the slower 6th grade math and reading/spelling sections. On the other hand, a number of the students were in the gifted and talented program. Ann noted that several of her students had test scores low enough to place them in the resource room for special assistance. Parental requests, however, had resulted in those students being retained in the regular 6th grade sections. Despite this diversity of skills, the class overall seemed fairly well balanced between high and low performers. Generally, students were eager to participate in classroom discussion and activities and were conscientious in completing homework and class assignments promptly. Overall, however, this class seemed to be the slower of the two sixth grade classes, and as Ann noted, was generally less academically proficient than her previous class at Adam.

The teacher, Ann Larson, had been in this elementary school for the past three years. She acknowledged that this environment was a sharp contrast to her previous experience in an inner city school in Portland. Ann, who was in her late 20's, maintained a well-ordered classroom. She firmly and

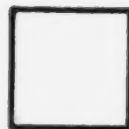
Teacher's  
Desk



X



X



X



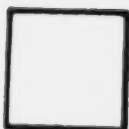
X



X



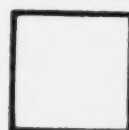
X



X



X



X



X



X



X



X



X



X



X



X



X



X



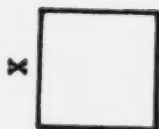
X



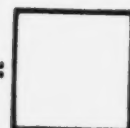
X



X



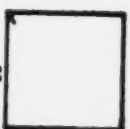
X



X



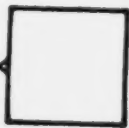
X



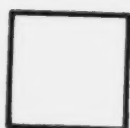
X



X



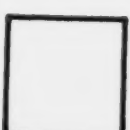
X



X



X



X

(Row used only for math class)

consistently monitored student talking and established clear expectations about students' homework responsibilities and on-task behavior. Students were assigned seats and appeared to have been originally grouped so that those needing more assistance were closer to Ann's desk. The diagram on the following page illustrates the seating arrangements for the class. This classroom setting as well as the tenor of the school further prompted attention to well-ordered activities. Since none of the classroom had doors and the two 6th grades opened into one another, it was important that the noise level remain manageable. Added to this physical feature, Ann acknowledged that she was most comfortable with a specific and consistent class structure, and noted that both 6th grades worked closely with the assigned texts and workbooks. Classroom activities involved a regular pattern of instruction, practice, questioning, correcting homework and discrete assignments, often specified in the text. Instruction was consistently organized and paced for the group as a whole rather than individualized.

The daily schedule for this class normally included the following activities and subjects:

9:00 - 9:05 am Homeroom. Students arrived in the class, had a few minutes to talk briefly with one another or their teacher and to organize themselves for the first class. Ann usually reviewed the day's schedule during homeroom and announced any changes or special activities.

9:05 - 10:00 am English. Ann began by asking students to clear desks of everything but the English assignment or text. During English, students regularly worked from the textbook on English usage, e.g., appropriate use of adjectives, adverbs and topic sentences. Ann introduced new material, asked students to read instructions orally, and gave them ample opportunity to practice or respond to sample questions in class. Classroom activity usually involved answering a section of questions from the text and/or correcting the previous day's assignments. Students regularly exchanged papers, corrected answers in class and read out grades based on a percentage scale which Ann recorded in the grade book.

10:00 - 11:00 Reading and Spelling. Students were grouped by ability for the reading section and moved into their respective 6th grade classrooms. Ann instructed the slower reading group which consisted of about 18 students. Class began with a brief period of silent reading, followed by one of a number of activities--spelling words and defining vocabulary words, oral reading, completing study sheets and correcting assignments.

11:00 - 11:30 am Social Studies. Students moved again into their homeroom classes. Social studies for Ann's class involved regular study units on Russia and on Africa and on the stock market (described in more detail below). Social studies usually included discussing current events, completing maps and worksheets, answering questions from the text and grading homework in class.

11:30 - 12:25 pm Math. Students changed classes for math. Ann instructed the high ability group which consisted of over 30 students. Activities regularly included correcting papers, reporting grades, doing sample problems, reviewing and answering questions and completing assignments.

12:25 - 1:05 pm Lunch and Recess.

1:05 - 1:35 pm Stock Market. Ann taught an innovative unit to her homeroom class as the second part of the social studies activity. During this class, students assumed the role of stock buyers and brokers and individually invested, tracked and summarized their costs and earnings from a \$10,000 initial investment in a series of stocks. Activities consisted of buying and selling stocks, and graphing and reporting the results of their activities.

1:35 - 2:10 pm Social Studies. During this period, Ann instructed the other 6th grade class in social studies. Activities paralleled those in her own social studies class.

From 2:15 pm on, students went to either music or PE followed by a study hall or optional band, choir or orchestra. Classes were dismissed at 3:30 pm.

I joined Ann's class as an observer on April 9th. During the following seven weeks, I attended class two days a week, usually on Tuesday and Wednesday although this schedule varied during the following weeks. I normally joined the class at 9:00 in the midst of the informal homeroom period, and stayed until students left for their elective classes (e.g., music and PE) at 2:15. As an observer, I collected several kinds of information that have been incorporated into this description. The most important has been a narrative description of all facets of the classroom assessment environment including--standardized and teacher developed assessments, assignments, teacher praise and reprimands, questioning strategies and instructional activities. In addition, I recorded interactions between teacher and students in six individual class periods, attending specifically to the students who volunteered, versus those who were called on, students who answered correctly versus those who did not. Several other valuable sources of information also supplemented these regular classroom observations. They included (1) informal discussions as well as more structured interviews with the teacher on her assessment practices and/or specific classroom activities, and (2) students' written commentary on what it takes to get good grades as a 6th grader. This latter activity provided perspective on the students' view about what was called for and rewarded in this sixth grade class.

The purpose of our classroom studies is to relate how teachers use assessments to diagnose, place, judge achievement, evaluate instruction and provide feedback to students and parents. In the following pages, I discuss the kinds of student characteristics assessed, the assessment strategies used most frequently in the classroom, the characteristics (e.g., teacher developed, objective, performance based) of these assessments; the methods of recording and providing feedback; and the assessment purposes relied on most frequently.



## Classroom Assessment Practices

In this classroom, the key assessment focus was evaluating students' achievement and mastery of subject content. Despite the emphasis on this activity, Ann also regularly assessed a number of other important student characteristics; namely, student aptitude, classroom behavior and social development and personality. Each played a role in the understanding of students, management of instruction and judgment of student proficiency.

### Student Characteristics Assessed

Achievement. Assessing students' mastery of subject content was a central activity, accomplished by quizzes, tests and the completing and correcting of daily assignments. Grading of students was based almost exclusively on outcomes from these three activities with considerable weight given to the completion of classroom assignments. All graded papers and tests were converted to a straight percentage scale and recorded in the grade book. Ann acquired almost all of these paper and pencil assessments from the publishers of the instructional materials. Few assignments or tests were exclusively teacher designed.

In addition to written assessments, Ann informally used oral questioning to track students understanding. Although class participation did not figure in actual grade practices, it did assist Ann in monitoring progress, adjusting the pace of instruction and in judging students' comprehension. Both oral questioning and written work (assignments and quizzes) assessed students' recall and understanding as well as their ability to apply information to other settings. Although assessment of higher order thinking skills was not a stated goal, Ann readily integrated this into much of her questioning. For example, in social studies discussions, she frequently asked students to compare and contrast or analyze information. Test items, in social studies in particular, did not have the same balance between recall of knowledge and terms and higher order thinking skills.

Student Aptitude. In addition to judging students' achievement, Ann employed a variety of methods to determine each student's general ability to handle the instructional material. Much of this occurred at the beginning of the school year as Ann noted students' responses to questions, skill in answering written questions and general facility in learning information. She noted that early "sizing-up" information had pointed out the need to slow the pace of instruction and to emphasize basic skills in English and reading for her current group of students. Knowledge about students' general ability also prompted her to dispute several planned 7th grade placements for her students. In one instance, for example, she disagreed with placing a student in a low math section because his in-class performance and level of questions and comments demonstrated abilities that were at a much higher skill level than indicated on the test.

Ann's overall judgment of student ability also contributed to her expectations about student performance. She consciously considered ability levels in grading certain borderline students and provided encouragement for effort to less proficient students. During recording of grades, she periodically praised students who were less academically proficient for showing improvement in scores, even though their scores were consistently lower than other students.

Classroom Behavior and Social Development. Ann clearly monitored students' behavior and their ability to manage themselves independently and in groups. She maintained consistent routines and clear rules on talking and inattentive behavior. The most frequent reprimands were for talking or disruptive behavior. Ann moved a number of students during my observations, locating them in settings where they could be more productive. She was also conscious of other general aspects of the student's personality, their social development, confidence, anxiety levels and self-concept, and discussed these characteristics when analyzing students' responses to classroom situations. Although those latter characteristics had no direct influence on grades, they did affect the way Ann managed and motivated the class.

Since assessing achievement levels was a major factor in this classroom, I would like to next examine the assessment strategies that assisted Ann in evaluating students' progress.

#### Assessment Strategies

During my seven weeks in the classroom, paper and pencil tests from the texts and district mandated tests, used to place and evaluate students' progress, played a major assessment role. During my 16 days of observation, students were given the following tests:

| <u>Test Type</u>         | <u>No.</u> | <u>Total</u> |
|--------------------------|------------|--------------|
| District Mandated        |            | 4            |
| Math Placement           | 1          |              |
| Levels Test-Math         | 1          |              |
| Levels Test-Reading      | 1          |              |
| Levels Test-English/LA   | 1          |              |
| Unit Test-Social Studies | 1          | 1            |
| Weekly Test-Reading      | 2          | 2            |
| Quizzes                  |            | 5            |
| English                  | 1          |              |
| Reading                  | 1          |              |
| Stock Market             | 2          |              |
| Math                     | 1          |              |
|                          |            | <u>12</u>    |

This section describes some of the characteristics and uses made of these tests including: (1) the test's purpose and origin; (2) the kind of questions included; (3) the feedback received by students; and (4) the use and interpretation of results.

### District Testing

During these weeks, district mandated tests absorbed a substantial amount of classroom time—four full classroom periods. Two specific types of tests were used (1) a district developed math placement test and (2) a district adopted test to measure student progress (the Portland Levels Test). In the district-wide math test, results from the 50-question test, developed by district teachers, determined student placement in (next year's) 7th grade accelerated class. As a followup, teachers received students' test scores as well as a list of students who would be placed in the accelerated class. Feedback on results did not, however, include information on how students performed on components of the test, or how the class overall performed on each math skill (e.g., fractions, multiplication). Students also did not receive specific score results.

Although test results did not give an indication of students' weaknesses and strengths, both students and teachers clearly viewed the test and its results as important. Ann, for example, was pleased at the significant number of students in her math class who were placed in the accelerated class. So test results provided an indirect evaluation of instruction for the teacher. A number of students, on the other hand, appeared particularly anxious during preparation for the test, asking a range of detailed questions about completion of the answer sheets. Ann noted, following the test that students had been uptight and participation in the accelerated program was important for students because of family expectations.

The Portland Levels Test was also administered during these weeks. The language arts, math and reading subtests took 3 full 55-minute time periods. The test's purpose is to track student growth over the year. However, this year the tests were being aligned with the local curriculum, and teachers felt that scores did not give an accurate reflection of student progress. Ann also noted, moreover, that the test results were not helpful in teaching. Nonetheless, these district-mandated tests involved careful preparation time, thorough attention to instructions and very specific attention to exact starting time. For example, Ann told students when it was 10 seconds before the test and asked them to begin exactly as the second hand passed 12. Ann also reminded students about strategies for taking timed tests and the importance of not wasting time on questions they were not sure of.

Ann received students' scores on the levels test for both spring and fall testing prior to the end of the school year. Although she reviewed the scores and noted students who showed gains and those whose scores declined, Ann used this information as a supplementary rather than primary information source about student progress. The results did not seem useful for assessing students or instruction. In addition, Ann did not mention how either the school or district intended to use this test information. In other words, these district-mandated testing activities, although obviously an important activity, seemed to have minimal effect on classroom practices or the diagnosis of students' needs.

## Classroom Testing

Ann's own testing activities combined with the grading of daily assignments and informal check of students' progress via questioning or observation were the central assessment features of this class. In respect to classroom tests, Ann noted that she relied almost exclusively on the tests provided in texts and accompanying teacher guides. In some instances, such as the end of unit social studies test, Ann supplemented these test questions with a few additional questions.

The tests and quizzes given students during these weeks were objective tests; questions were multiple choice, short answer, and true-false. The social studies unit test also included a brief essay question. Prior to a major test, Ann carried out a thorough review with students and described exactly the kinds of questions to expect on the test. During the review for the social studies test, the only major unit test given during my observation, she indicated that it would include a brief essay at least 4 sentences long and spelling would be counted throughout. Both review and test, in this instance, emphasized factual information--knowledge of terms and recall--to a greater degree than had been apparent in the classroom discussions. At the conclusion of this test, Ann promptly reviewed answers, returned graded papers on the following day and further discussed answers and reviewed questions that posed particular problems. Because students had unexpected difficulty on this test, Ann modified the grading and decided to grade on the "curve" rather than use the usual percentage scale and to not count spelling errors in the essay question. Her review of students' grades and informal questioning about the difficulty of individual test questions, prompted Ann to conclude that the lower grades were due to (1) students' lack of familiarity with multiple choice questions that called for all correct answers and (2) insufficient studying by students. She also noted that students had performed far better on homework assignments than on the test, despite the fact that she viewed the test as relatively easy. This unit test was the most important classroom test given students during my observation.

## Classroom Assignments and Diagnosis

But tests were definitely not the only way Ann judged students' progress in a subject. Assignments, completed almost daily and corrected in class, provided an equally important source of information and contributed significantly to a student's overall grade. For example, in the mid-term progress report, 71 percent of the 2100 possible points (20 assignments) were based on assignment grades. Assignment results were key influencers of student grades. Similar to the test questions, assignments were usually part of the curriculum materials and frequently called on students to answer a series of questions. Almost all assignments were graded; in fact, students frequently asked "Will this be graded?". Assignments were regularly corrected in class and scores were converted to a percentage scale which the student read out loud to Ann for entry into the grade book.



Although Ann frequently collected papers after grading, she also used other informal methods of tracking students' general progress. To maintain a general sense of the group's progress, she regularly asked students to raise their hands and indicate how many questions each had missed. In-class assignments were often preceded by practice assignments of one or two questions or problems. While students completed these questions, Ann consistently circulated throughout the room observing each student's progress. The number of questions raised, the speed with which students completed the practice questions all indicated how much additional clarification needed to be provided.

In this setting, the prompt completion and correcting of assignments played as much, if not more, of a role in the grading process of students as did tests. For students it also meant that their performance was constantly visible before classmates. The process of recording assignment scores was also a time when Ann verbally acknowledged good grades. For the skillful student it was a chance for others to hear how well they had scored and to be verbally praised for the performance. Other students, who repeatedly had to acknowledge low grades, must have viewed the daily display with apprehension. While the regular grading and verbal reporting probably motivated some students to complete assignments and to stay current with their classroom work, the daily recording of grades also prompted students to constantly compare themselves to their peers.

#### Evaluation Criteria

Completing assignments on time was an important criteria for success. What other evaluative criteria other than assignment completion were specifically noted by the teacher in preparing students for classroom work?

Although the overall criteria for performance may be discussed early on in the school year, I looked for explicit comments to students about what criteria were being evaluated. During my observations, I recorded five times when Ann clearly noted criteria for her assessments. Most of these involved evaluating some type of written product—the essay on a test, a descriptive paragraph, the stock market booklet. Because these criteria tell students explicitly what assessment expectations are, I've outlined the criteria in the following:

- |                          |                                                                                                                                                                          |
|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| a. Test essay            | Write a paragraph with at least four sentences. Spelling counts. The paragraph should be well thought out.                                                               |
| b. Descriptive paragraph | Spelling counts. Use at least 10 sentences. Sentences must all support a topic sentence. Make certain sentences make sense. Use descriptive adjectives and action verbs. |
| c. Revising paragraph    | Watch spelling and use a paragraph format. Link sentences to topic.                                                                                                      |

- |                         |                                                                                                                  |
|-------------------------|------------------------------------------------------------------------------------------------------------------|
| d. Stock market booklet | Neatly organized, accurate descriptions. Use paragraphs; expect clear, understandable sentences; should be neat. |
| e. Reading stories      | Use nice, clear, loud voice.                                                                                     |

In several instances, Ann used a student's assignment to illustrate her standards. For example, she read one of the descriptive paragraphs assigned in English to the class, noting how effectively it was written. Most evaluative criteria were communicated verbally rather than written down, and most focused on straight-forward characteristics such as spelling, number of sentences in a paragraph, or clarity of the sentences. During my observation, students had only one opportunity to revise an assignment and improve the quality of it. For most students, this involved making the assignment neater rather than improving content.

### Students' View

What did students perceive as the important criteria for success in this 6th grade classroom? To answer this, we asked students to write a letter to an imaginary 5th grader describing what it takes to be successful in 6th grade. Students identified the following issues as most essential.

|                                                           | <u>No. of Responses</u> |
|-----------------------------------------------------------|-------------------------|
| o Get homework done and turn it in on time                | 20                      |
| o Listen, pay attention, don't talk and follow directions | 15                      |
| o Study notes and past assignments for tests              | 15                      |

The next most frequently mentioned responses of students were (1) study and try hard (9 responses), (2) check assignments for mistakes, and (3) be neat, use nice handwriting and correct headings (8 responses for both). According to these responses, assignment completion and appropriate behavior were the criteria students recognized as most significant contributors to good grades. Surprisingly, few students mentioned quality of work or participation in class. A summary of student responses is included at the end of this paper.

### Questioning Strategies

Daily questioning activities allow teachers to informally monitor students' understanding, both individually and as a group. To analyze teacher-student interactions in this class, I recorded question and response

patterns in three English and three reading/spelling class periods and noted students who

1. volunteered responses and answered correctly or incorrectly
2. were called upon and answered correctly or incorrectly
3. were asked to read orally
4. initiated questions and
5. received praise for their participation.

These interactions are summarized in detail in the following tables.

In reviewing responses, these patterns were clear: (1) students eagerly volunteered to answer questions and most volunteers answered correctly; (2) relatively few students were called upon to answer questions; (3) students spent a significant amount of time asking questions about assignments or about correcting papers; and (4) students who volunteered and participated actively in class were far more likely to receive praise than other students.

Many students volunteered to answer questions. In English, for example, 68 percent of all questions asked were responded to by volunteers. In reading/spelling, 82 percent of the questions were answered by volunteers. As the table on the following page illustrates, most volunteers also answered correctly--in English, 13 of 17 responses (76%), in reading/spelling\*, 21 of 24 (88%). The number of students called upon in each class was significantly lower. In English, students were "called upon" to answer 32 percent of all questions; all responded correctly, and almost all were high performers. In addition, 11 English students (42 percent of the class) did not participate, during these class periods, in question answering activities. In reading/spelling, 17 percent of the questions were directed at specific students; these students, as the table illustrates, were much more likely to respond incorrectly than correctly.

Since Ann relied on volunteers to monitor student's understanding, some students could readily slip by without being called upon and without attending to classroom discussion.

Students' Responses to Questions

|                  | N  | V   | Ⓥ   | C   | Ⓢ** |
|------------------|----|-----|-----|-----|-----|
| English          | 26 | 16% | 52% | 0   | 32% |
| Reading/spelling | 11 | 10% | 72% | 14% | 3%  |

\* I recorded only Ann's homeroom students in the reading and spelling class. The figures recorded reflect the responses and participation levels of 11 students, not all students in the class.

\*\* Symbols represent the following: V = volunteer, answer incorrectly; Ⓥ = volunteer, answer correctly; C = called on, answer incorrectly; Ⓢ = called on, answer correctly; A = student asks question; R = student reads orally; P = student receives praise; N = number of students.

Students not only answered questions eagerly--a sign of their involvement--they also asked a great many questions themselves. In fact during the English sections over three times as many questions were asked by students as by the teacher. In reading, approximately two times as many questions were asked. The vast majority of these questions had to do with students correcting papers in class and with requesting help on assignments. The somewhat surprising number of questions from students about management rather than content issues illustrates the time and effort devoted to the grading of papers, particularly in determining how many points an answer should receive.

Finally, Ann consistently provided praise for volunteering and answering questions correctly, for reading orally and for reporting high grades. Consequently students who participated actively in class--volunteered frequently, etc.--were far more likely to be verbally praised and reinforced. There seemed to be far fewer opportunities for reinforcement of students who were less proficient academically or verbally, and less reinforcement occurring in the reading/spelling than in the English.

Summary Tables\*\*

|   | English (n=26)  |                            |       |       | Reading/Spelling (n=11) |                            |       |       |
|---|-----------------|----------------------------|-------|-------|-------------------------|----------------------------|-------|-------|
|   | Total Responses | No. of Students Responding |       |       | Total Responses         | No. of Students Responding |       |       |
|   |                 | Boys                       | Girls | Total |                         | Boys                       | Girls | Total |
| V | 4               | 2                          | 1     | 3     | 3                       | 1                          | 2     | 3     |
| Ⓥ | 13              | 4                          | 6     | 10    | 21                      | 3                          | 6     | 9     |
| C | 0               | 0                          | 0     | 0     | 4                       | 2                          | 1     | 3     |
| Ⓒ | 8               | 1                          | 5     | 6     | 1                       | 0                          | 1     | 1     |
| A | 83              | 13                         | 10    | 23    | 56                      | 4                          | 7     | 11    |
| R | 23              | 10                         | 8     | 18    | 7                       | 2                          | 5     | 7     |
| P | 37              | 12                         | 8     | 20    | 7                       | 2                          | 5     | 7     |

Providing Feedback to Students. In this class, grades recorded in the gradebook were the major record of student progress. Both 6th grade teachers used a computer to record and average grades from assignments, tests and quizzes. At midterm, students received a printout of their grades on all activities and their grade average so they could identify exactly what assignments were missing or had posed problems for them. Ann relied primarily on this numerical summary to give feedback to students on their progress. She also acknowledged that final grades in borderline cases were affected by the amount of effort students displayed. There were no other permanent records such as a file of sample papers or anecdotal records maintained on students. Students received a grade report with letter grades at the end of each term.

\*\* Symbols represent the following: V = volunteer, answer incorrectly; Ⓥ = volunteer, answer correctly; C = called on, answer incorrectly; Ⓒ = called on, answer correctly; A = student asks question; R = student reads orally; P = student receives praise.



In addition to the feedback provided by grades, Ann provided a considerable amount of informal verbal praise to students in class. Praise such as "good" or "very good" was most likely to occur in these instances: (1) after a student had responded correctly to a question, (2) read a section in the reading book, (3) reported a good grade on an assignment, or (4) asked a relevant question. As a group, students were praised for appropriate behavior such as beginning classwork promptly, or keeping noise to a minimum. Praise also was used to acknowledge and encourage effort. For example, during reading activities, students who had difficulty with oral reading frequently received a comment of "good" or "very good." Also, students who volunteered or who tended to be disruptive frequently received praise for their contributions and class participation. Quieter students, who volunteered infrequently, were less likely to receive praise. Ann, also, consistently provided positive comments to students as they reported grades on assignments. Students with high grades regularly received "good" or "very good." In a number of instances, students who normally received lower grades received praise for evidence of improved work even though their grade might have been average or lower.

Although students corrected their own and other students papers almost daily, students did not have any personal involvement in evaluating themselves. For example, they did not keep track of their grades on assignments, nor keep a folder of writing papers nor complete any rating scales or checklists critiquing their projects or those of their peers. As may be true in most classrooms, grading and evaluation was primarily an external activity and students had little opportunity to apply evaluative criteria to their own work or work of peers.

#### Purposes for Assessment

Although grading, including evaluating tests and assignments, was the most important assessment activity in this class, judging students' achievement was not the only assessment activity. For example, Ann also used assessment to diagnose group and individual needs, to place or group students, and to motivate and control. This section describes how these types of assessments were used.

Diagnosis. Assessing the progress and needs of students, individually or as a group, affects the pace and character of instruction. As noted previously, diagnosis of the students' progress as a group occurred frequently and involved a myriad of activities—informally checking papers as students completed assignments, asking questions of students with varying skill levels, observing students' progress on practice problems, and monitoring the number of questions asked about an assignment. Ann also noted that she used pretests with her math class to see if a unit should be taught or skipped. In several instances, diagnosis clearly led Ann to reteach a lesson or review certain concepts, but it was not linked to individualizing learning activities. Instruction was provided on a group basis. Students who needed extra help were asked to come in after class for assistance. Ann arranged small group assistance two or three times during the time I observed, at other times she expected students to contact her if they were having difficulty.

Although Ann was aware of several students with potential learning problems, these problems were primarily managed in the context of regular group instruction. For example, one student consistently reported low grades on assignments and tests. Yet he was a fluent reader, frequently answered questions in class correctly and received average scores on the district testing. In class, Ann called on this student frequently, consistently praised him for his efforts and reprimanded him for incomplete assignments. Despite these strategies, the student continued to perform poorly. Although Ann modified group instructional activities to assist this student, a more individualized diagnosis of his specific problems seemed to be required.

Placement and Grouping. In this school, students received placement tests in reading and math at the onset of each year. The placement tests, from Harcourt Brace, consisted of a series of reading selections and comprehension questions and math problems covering addition and subtraction, rounding numbers, multiplication, division, fractions, decimals and geometry. New students also took these two tests, and classes were assigned based on test results. According to Ann, two other factors played into modifying final decisions about placement—teachers' observations of students during the initial week of class and parental requests. Ann noted that some of her reading students, whose test scores would have placed them in the resource room, had been placed in her room by parental request. Parents wanted to make certain their students were progressing at grade level and not segregated into a special learning context. This resulted in a reading class of substantially differing abilities, but Ann felt all were managing to keep up with grade level work.

The initial weeks of class also provide time for teachers to "size up" the class as a whole, as well as individuals in it. Ann relied on a range of standard classroom activities—question answering, completing assignments, classroom discussions—to help her determine the groups needs and abilities. During these initial days, she noted that the students in this class were definitely not as proficient as her last year's class and decided that she would place more emphasis on basic skills and limit the amount of writing assignments students were expected to complete. Initial "sizing up" activities also prompted Ann to consider eliminating the unit on the stock market because she was uncertain about the students' ability to handle it. In retrospect, after seeing students' writing assignments and their success with the stock market unit, Ann indicated that these modifications were probably not needed and questioned the limited amount of writing that she required of students.

Student Motivation. Ann frequently used tests and assignments to motivate and control student behavior. For example, in introducing activities for the week, she regularly told students that they would be tested on the material on Friday. Ann also kept students focused on classroom activities by telling students that papers would be corrected at the end of class. The importance of an assignment was determined by whether or not it was graded. For example, when explaining the required booklet on stock market transactions, Ann specified at the onset "This will be graded." Students, also, frequently asked "Will we grade or correct this?" Ungraded practice papers were often treated as unimportant; Ann frequently reprimanded students for crumbling them

up and throwing them in the wastebasket. Correcting and grading papers in class and completing homework which would be graded the following day was a routine part of the daily activity and a method of keeping students constantly aware of and progressing toward their next tasks.

The only subject this pattern did not occur in was the stock market activity. During this social studies class, students reported on a personal (imaginary) investment activity. Rather than a series of short term assignments, corrected on a daily basis, students engaged in a long term activity and summarized the results in a final booklet. The students' focus was on learning how to complete and clearly describe their investment activities. Although students did not have daily written assignments, the vast majority stayed motivated and engaged throughout the activity and responded positively to the departure from typical class patterns.

Evaluating Instruction. A teacher's ongoing instructional evaluation is primarily an internal process—a daily review of how attentive students were, how quickly they grasped material, or whether concepts needed to be retaught. However, a more visible example of evaluation occurs when students perform poorly on a major test. For example, in evaluating results of the social studies test, Ann informally considered test difficulty; format of the questions; time spent in review; and students' general results on homework. She did not, however, consider whether the test questions themselves were appropriately linked to instruction; whether questions tested what was taught; whether the test judged students' actual understanding of material. Ann, instead concentrated on dealing with grading issues (e.g., grading on the "curve" rather than with the normal percentage scale). In this instance, the test seemed to have little value as a means of judging or modifying instruction and there seemed to be relatively little link between testing and instruction.

Feedback to Students and Parents. Grades from assignments and tests were the major feedback provided students and parents. These scores were reported in detail on the midterm computer printout which provided a summary of total points available, points earned by the student on each assignment, and an average of the number of points earned. This average was indicated in percentage points and equated to a letter grade (e.g., 90-100 = A). In addition, Ann indicated on the printout whether the student's work was below average, average or above average. The printout did not include any summary comments about students' attitudes, progress or behavior. Ann noted that the computerized reporting was a new system, one that alerted parents to their son or daughter's progress well before grades were released. Both sixth grade teachers felt it provided important documentation of a student's daily work and helped circumvent potential conflict with parents about grades. Parents and students also received regular report cards at the end of the term. Grades were computed on a percentage scale: A = 90-100, B = 80-90, C = 70-80, D = 60-70, F = below 60. The scale was used in all subject areas.

As I mentioned earlier, students also consistently received feedback on both performance and in-class behavior. Ann provided regular reinforcement for homework grades, test results, answering questions in class, reading

orally, and initiating questions. Periodically she encouraged those who missed questions by saying "good try." Ann also provided feedback to the group and individuals about on-task behavior and closely monitored disruptive students and general class behavior. She praised positive group efforts, commented on productive work habits, and firmly disciplined students who disrupted the class or failed to hand in homework on time.

Students' classroom work was judged in terms of being right or wrong and scored according to the percent correct. For assignments that required a different kind of evaluation such as the stock market booklet, or a descriptive paragraph, Ann gave clear, specific criteria. The criteria, however, emphasized characteristics that were more easily measured--spelling errors, length of paragraph, number of descriptive adjectives rather than content issues. In the regular subject areas, she concentrated on more measurable and objective characteristics and on assignments that could be easily and efficiently corrected.

Since most classroom feedback was tied to grades or classroom participation, students who generally performed well in those areas received the majority of positive feedback. For the quieter and less academically proficient student there were fewer opportunities for verbal reward and seemingly less opportunity to feel successful. In addition, the kind of feedback provided did not seem to help students learn to perform more effectively. For example, although Ann suspected that students had not studied adequately for the social studies test, I did not observe her specifically discussing how they might review more effectively for the next test. Feedback in this setting seemed to provide more consistent reinforcement for competent learners, and less assistance or benefit to those who were less academically skilled.



## Summary Comments

At the onset of this chapter, I noted that the objective of this study was to further understand classroom assessment practices and describe how assessment is used to diagnose, place, evaluate achievement and instruction, and provide feedback to students and parents. In summarizing my observations in this 6th grade class, these points seem particularly relevant.

Assessing student achievement by grading is the major assessment activity in this classroom. This assessment occurs daily, in every subject area, through the scoring of classroom assignments and the grading of quizzes and tests. In fact, assignments are just like mini-tests and overall have as much if not more weight on a given grade than the tests themselves. In evaluating assignments, the major criteria are completeness and correctness. Scores on homework were recorded daily in almost every subject area.

In this classroom, assignments as well as tests were almost always part of the instructional materials for a unit. Ann depended upon the publishers of the curriculum materials to provide relevant, well-written assessments. In only a few instances did she supplement the test with additional questions.

Both assignments and tests were paper and pencil assessments usually calling on students to answer multiple choice, short answer and fill in the blank questions. Students had very few assignments that required the teacher to judge products or performance apart from homework. The primary exception to this were students' activities in the stock market unit.

Although district-mandated testing absorbed a fair amount of instructional time, this teacher depended exclusively on her own tests and assignments to judge students' progress. The formal tests had little or no influence on classroom activities and were used primarily to facilitate student grouping for the following year.

Students received frequent praise and reinforcement for their grades, participation in class discussions and responsible behavior. Academically proficient students, however, tended to receive and have opportunity to benefit more from the reinforcement available than did students who were less accomplished in the classroom.

Student Comments--How to Get Good Grades

| <u>Comments</u>                                                                   | <u>No. of Responses</u> |
|-----------------------------------------------------------------------------------|-------------------------|
| Get homework done on time, turn it in on time                                     | 20                      |
| Listen, pay attention, don't talk, follow directions                              | 15                      |
| Study notes and past assignments for tests                                        | 15                      |
| Study and try hard                                                                | 9                       |
| Check assignments for mistakes, including misspelled words, punctuation, capitals | 8                       |
| Be neat, use nice handwriting and correct headings                                | 8                       |
| Understand what you're learning, ask if you don't                                 | 5                       |
| Be good, don't get in fights, don't talk back                                     | 5                       |
| Take time and do a good job                                                       | 3                       |
| Concentrate on work, don't goof off                                               | 3                       |
| Try to do all extra credit and extra reading                                      | 3                       |
| Complete all assignments, complete make-up                                        | 3                       |
| Study in a quiet place, right after school                                        | 3                       |
| Help teacher, get on her good side, be nice to her                                | 3                       |
| Never talk back to teacher, don't mess with her stuff, or get out of seat         | 3                       |
| Answer questions when called on                                                   | 2                       |
| Never cheat or copy answers                                                       | 2                       |
| Read thoroughly, don't skim, following along when others read                     | 2                       |

Student Comments--How to Get Good Grades

| <u>Comments</u>                                                 | <u>No. of Responses</u> |
|-----------------------------------------------------------------|-------------------------|
| Make up bad grades                                              | 1                       |
| Don't tattle or make fun of others                              | 1                       |
| Don't complain about long assignments                           | 1                       |
| Consequences/warnings                                           |                         |
| If homework is late, you'll write<br>25 sentences               | 1                       |
| Get a 0 if assignments not completed                            | 1                       |
| Get stricter at end of year                                     | 1                       |
| Extra hints                                                     |                         |
| Get sleep and eat a good breakfast<br>before tests              | 1                       |
| Write assignments down, check off<br>when completed             | 1                       |
| Sometimes teachers give answers<br>by accident                  | 1                       |
| Good notes make it easier to study,<br>it helps remember things | 1                       |
| Good grades are important                                       | 1                       |

CASE STUDY #3  
Methodology

This report is based on over two months' participant-observation in the sixth grade classroom at Mount Stanton School (1). In the period from 1 April to 3 June, 1985, this researcher spent an average of 15 hours per week with the students and teachers of the sixth grade, observing the instructional process in the classroom, assisting with oral reading groups, doing some one-on-one work with students working on individual projects, following the students as they went off to their other classes, occasionally joining in the Lower School faculty in the lunch room and their after-school meetings, and informally interviewing students, teachers, and staff.

Further, a variety of documents relating to classroom assessment were collected: samples of homework and classwork assignments in math and language arts; the class newsletter and various notices to parents; evaluations by camping trip counsellors; the semi-serious personal awards list from the class camping trip; student self-evaluations of their year's work; student-generated lists of ways to succeed in the sixth for incoming sixth graders; and a sample of mid-year and all available year-end student evaluations. In addition, for the single textbook unit used in the class, an analysis of the cognitive domain of questions was conducted.

In the paragraphs below I will attempt to characterize evaluation and assessment in the Mount Stanton sixth grade. Assessment is an ongoing, visible part of the instructional activity in the sixth grade classroom. Student-teacher interactions were documented in fieldnotes; sample homework and classwork assessments supplement these observations.

The two sixth grade teachers devote many hours to constructing prose evaluations of each student twice yearly. Final evaluations were available for 38 of the 42 sixth graders; a sample of 12 mid-year evaluations was also collected. These evaluations--especially the final evaluations, which report the period during which I was present in the classroom--provide direct insight into which aspects of academic and personal behavior are assessed and how they are weighted. Further, when compared and contrasted with classroom

---

(1) The school name, as well as the names of all personnel and students are pseudonymous. I wish to extend my thanks to the staff and students of the Lower School for their generous cooperation with and support of this research. Most gratefully I acknowledge the limitless interest, attention, and patience of the sixth grade teachers. They maintained interest despite all my intrusions and gave fully of their intelligence and broad experience to this project. The students of the sixth I especially thank for their time and curiosity. They made me feel welcome, sometimes even useful--a rare and gratifying experience for a researcher.



observations, the teachers' final evaluations reveal the degree of congruence between informal and ongoing performance assessment and formal, permanent assessment records.

The materials available from Mount Stanton also offer some insights in to the students' view of assessment. Two student assignments, their year-end self-evaluation and their lists of ways to succeed in the sixth, are compared with the teachers' final evaluations. Self-evaluations were available for 18 students. For 11 students these self-evaluations are substantial enough that they can be compared on an individual, as well as group, basis with the teachers' final evaluation report. Lists of ways to succeed in sixth, numbering from one to 12 suggestions, were completed by 22 students. While the comparison of self-evaluations and teacher evaluations primarily addresses specific aspects of academic and personal growth, the ways to succeed lists suggest how closely students' and teachers' understandings of desirable classroom behavior conform.

In order to create a context in which the assessment activities of the Mount Stanton sixth grade can be properly understood, the first major section below will provide background information on the school and the sixth grade itself. The structure and atmosphere of the sixth grade will be described in some detail, since the instructional method and philosophy of its teachers determine, to a great extent, the form and function of the assessment activities to be reviewed.

The succeeding sections will address specific assessment questions:

- What are the forms in which assessments are recorded, both in informal teacher records and school files?
- What are the sixth grade teachers' principles for assessment, i.e. what is the range of student characteristics and behaviors that should be assessed?
- What are these teachers' strategies for assessment; how are assessments of student performance actually conducted?
- What are the instructional or other pedagogical purposes that assessment serves in this classroom?
- How and how effectively are assessment principles communicated to the students; in what ways do students participate as co-active members of a cooperative assessment process?

In order to address these issues, especially the final question of feedback to students, the various assessment documents described above will be compared and contrasted. The study focusses almost exclusively on the activities of the sixth grade as a homeroom group, thus the subjects of math, language arts, and reading, for which the homeroom teachers are responsible, constitute the primary academic foci of the report.

## "Running The Sixth" at Mount Stanton School

### The School

Mount Stanton School is a well established, private, nonsectarian, independent, coeducational preschool-12 academy located on a wooded campus in a major West Coast metropolitan area. It enrolls 575 students, ten percent of whom receive partial or complete waiver of the \$5,300 annual tuition fee. Students are drawn from the city, its suburbs, and surrounding semi-rural communities, primarily the children of professionals. Parents must provide transportation to school. Their level of involvement in their children's education is generally high.

The school has four divisions, each housed in its own building: a preschool enrolling children from age three-and-a-half through kindergarten; the Lower School, of which the sixth grade is the final class; the Middle school, grades seven and eight; and the Upper School, grades nine through twelve. Enrollment per grade is smallest in the Lower School, averaging 40 per grade, and greatest in the Upper School, averaging 50 per grade. In addition to the divisional buildings, the campus boasts a classroom and studio buildings, a cafeteria/assembly building, an auditorium, a wood shop, gymnasiums and playing fields, and an administration building. Each division of the school has a principal; general administration and planning is directed by a headmaster who reports to the corporate board of trustees.

The greatest transition for students at Mount Stanton takes place after the sixth grade, when they shift from a single class per grade to achievement-grouping for their various subjects. The transition from the Lower School also marks a transition toward the Upper School's more traditionally structured, college-preparatory curriculum. During the Middle School years the students also move from the ungraded evaluation system of the Lower School toward a graded system, fully instituted by the ninth grade. In grades seven and eight prose evaluations are still the standard reporting form, however grades are also computed and are available upon request of student or parent. Thus, the sixth grade is a pivotal year for Mount Stanton students; the evaluations and assessments of the sixth grade teachers are informed by this impending change in student life.

In the Lower School teachers enjoy a high degree of discretion in their curriculum and pedagogy. However, the school has a well articulated philosophy for learning and teaching. The teachers' primary purpose, according to the Lower School Parent Handbook is:

...to help children develop intellectually, socially, and physically, and, in the process, through self-discipline and through the satisfaction derived from doing a task well, become self-reliant individuals.

Throughout the Lower School each grade is maintained as a single, team-taught class. Students break out of their contained classrooms into smaller groups for art, music, shop, science, and French. Physical education is a daily class, conducted as a coed activity in grades one through five, but separately for boys and girls in sixth.

Mount Stanton places considerable stress on character development, as well as academic achievement. Character development is accomplished in part through experiential learning. Educational outings are frequent, ranging, for the sixth grade, from a walking tour of the city to a visit with the state governor, from trips to soccer tournaments with out-of-town schools to a week-long naturalist and camping trip to the ocean. These excursions are tied to classroom learning, in the sixth requiring advance discussion, reading, and/or film-viewing, as well as individual writing and drawing projects.

### The Sixth

The students in the sixth grade are "run" (in teacher jargon) by a coed team of five years' standing. Clint is the Lower School's most senior teacher, the assistant principal and a Mount Stanton sixth grade veteran for over 20 years. He and his principal describe him as the "creative, imaginative" partner and, indeed his classroom style is dramatic, engaging, and affectionate. Clint is a veritable legend at the school; students from years ago, as well as current Upper School students, make time to drop in on his class. Brenda is, both in self-description and according to her partner, the "organizational mind", the one "who makes things run." She began at the school five years ago, having found public school teaching too regimented and fraught with disciplinary problems. Brenda projects a low key, sincere, and thoughtful personality to the students; they know she is a caring person, interested in any issue they bring to her.

The two teachers divide some of the labor of running the sixth. Brenda maintains meticulous class records, corrects the majority of the homework and classwork, and generally keeps track of time, people, and things. Clint takes on much of the discipline work, teaches one of the PE classes, and helps supervise the students when they go to shop class. Clint's extra efforts provide some free moments for Brenda to work on the students' papers and undertake administrative work, such as drafting the evaluations.

The atmosphere of the classroom is highly informal, even to the point of giving the appearance of being unstructured. In fact the informality is coupled with a highly structured system made as explicit as possible to all participants. One of the teachers' primary tasks in the fall, they report, is to make the system for running the sixth clear to the students. All the members of the sixth are on a first-name basis. The teachers are warm and physically, as well as verbally, affectionate toward the students. Students feel free to bring any concern or anecdote, school or nonschool, to either teacher and receive a receptive, interested hearing. Brenda and Clint tolerate relatively free movement in the classroom and the students can generally come and go as they need to, provided there is no clearly

group-focused activity going on in which they have a responsibility to participate. Considerable student-student interaction is allowed around the general instructional activity, in fact the seating arrangement (see Figure One) is designed to facilitate this. Queried about this student-student, rather than teacher-student seating orientation, Clint explained that:

Kid-kid interaction is as important as what goes on between kid and teacher....If you don't encourage that, you miss out on so much. They learn from each other. They get skilled....It (talking to other kids about instructional material) puts the capper on some skills.

The 1984-85 sixth grade was large, 42 students. Sexes were fairly evenly divided, 23 boys and 19 girls. Brenda and Clint feel the class runs best when there is near parity in the sex of the students. Eleven of these were new to Mount Stanton this year, close to the usual 25 percent new students they encounter every fall. Students are placed into the school at this age for basically two reasons, according to Clint. On the one hand, there are children who have begun to have difficulty in the public school, causing disciplinary problems and/or becoming bored and unmotivated. These parents hope their children will profit from the more individualized attention available at Mount Stanton. On the other hand, some parents feel that their children, while performing and behaving well, are not realizing their full potential. They look to the stimulation the wide range of activities and projects for student research offered by Mount Stanton, as well as its emphasis on individual development. Five or six of the students were preparing to leave Mount Stanton after the sixth grade, transferring to their local public middle school. This is the usual level of attrition; some parents view the Lower School experience at Mount Stanton as the crucial period in structuring their children's attitudes toward learning. All these students were very reluctant to leave Mount Stanton; some had serious anxieties about the graded and more hierarchically structured environment they expected to enter.

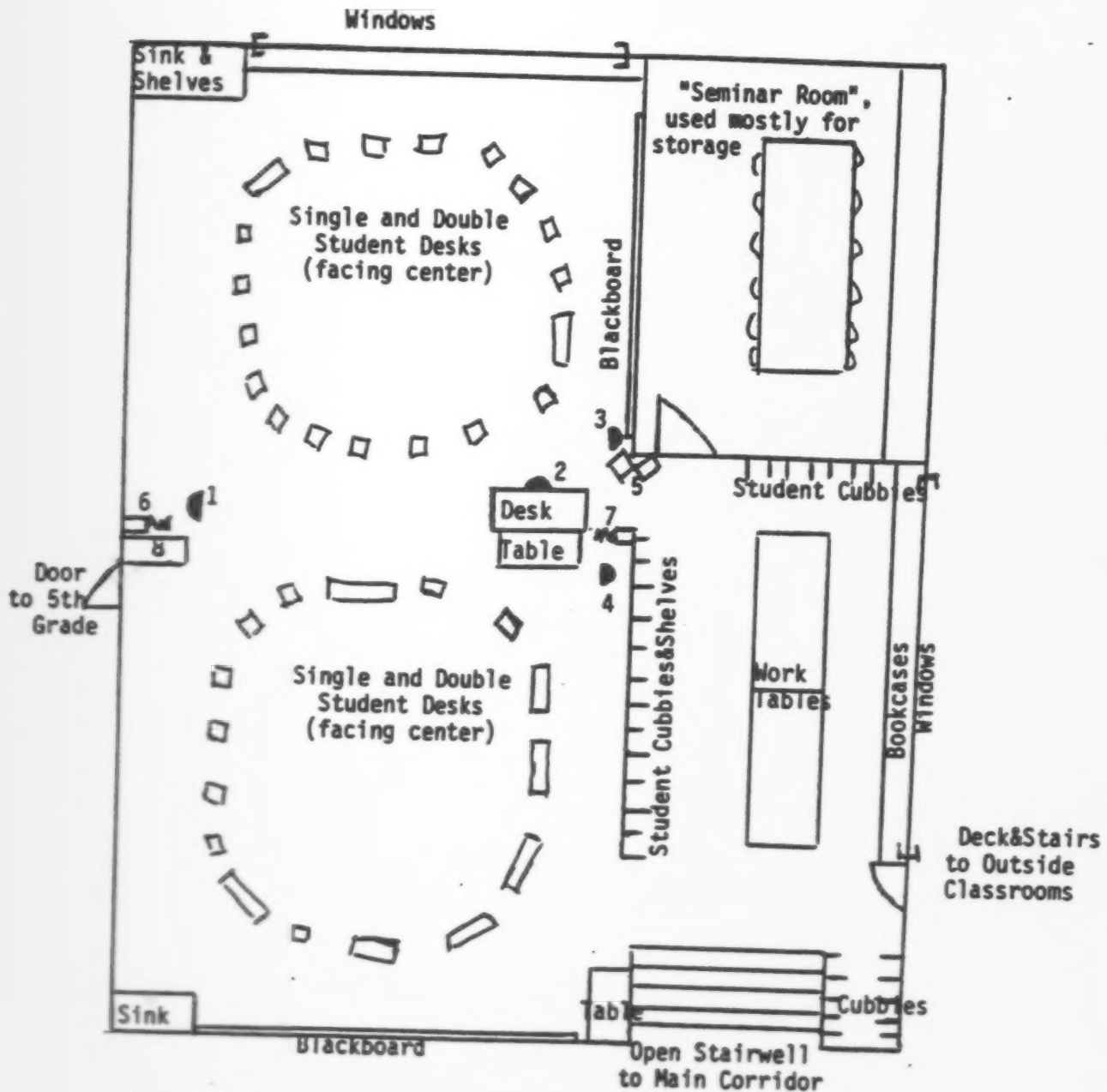
Ages of the classmates ranged from 10 to 13 years. Several students who had come up through Mount Stanton from preschool had entered the Lower School early and thus were ahead of their age group. Nor is it uncommon for students to repeat a year at the school, leading to older-than-average sixth graders. (One student was a repeater from the previous year's sixth. Noam had not performed well academically or personally and he had agreed with his parents and teachers that another chance to do the sixth would be a good experience.) The students also ranged widely in maturation, some well into puberty and others still very much children.

While not all the students in the sixth were above average in either ability or achievement, the school's overall student body ranks above the eighty-fifth percentile in nationally normed tests. Decisions to request repeating a grade contribute to this, as well as refusal to keep students who are not suited for the responsibilities that the school's students are expected to fulfill. Brenda and Clint related a case from this fall in which



Figure One

Physical Plan, Mount Stanton School  
Sixth Grade Classroom



1. Teachers' Speaking Chair
2. Teachers' Desk (Brenda's work area); Homework "in" box on corner
3. Extra Teachers' Chair (researcher's station)
4. Visitor Chair (researcher's station)
5. Student Assignment Files
- 6 & 7. Folding Floor-to-Ceiling Partition
8. Storage Cabinet

they had, in fact, refused to run a student who was academically and emotionally unprepared for their sixth grade. He was combative, did not punctually meet his classes, and performed far below their minimal reading levels. They were on the verge of requesting his removal when he was caught stealing in the classroom and was summarily expelled by the principal. The episode led to a soul-searching faculty meeting in which other teachers admitted they had passed him on because he was simply too disruptive. New guidelines for handling, and dismissing, problem students were established as a result.

Clint and Brenda structure their curriculum around a series of major topics, many of them traditions for the sixth that Lower School pupils look forward to for many years. Each successive topic requires a student project that, by the end of the year, entails library research, several drafts of a written report, an oral presentation to teachers and classmates, and a wall mural depicting aspects of their topic and that serves as an aid in their presentation. A first topic in the fall is used as an icebreaker, to help integrate the new and continuing Mount Stanton students, and to facilitate Brenda and Clint's getting to know the students well. Each is assigned a partner and required to write a biography of that classmate, introduce him/her to the class, and draw a portrait. Subsequent topic foci included a biography of an historical personage, a marine science or history study, and a report on the city in whose suburb the school is located.

The sixth makes almost no use of textbooks. The single exception is Scott Foresman's IMAGE I (1977), a short story and essay collection, out of which one unit is assigned, taking up several weeks of reading time just before mid-year. Reading is taught orally, in small reading circles. The students choose among a variety of offered books, novels and non-fiction essays. Selections late this year included Farley Mowatt's THE DOG WHO WOULDN'T BE and OWLS IN THE FAMILY, Esther Hautzig's THE ENDLESS STEPPE, and Irene Hunt's ACROSS FIVE APRILS. Language arts is taught through the major project work and short, in-class and take-home assignments created by Brenda and Clint and dittoed worksheets drawn from a variety of sources. Math, too, is conducted without a textbook. Clint teaches math from notes and assigns problems of his own devising and worksheets from various published sources for homework.

Much of the interaction in the sixth is one-on-one between teacher and student. There is relatively little group rhetorical question-answer work. Rather, in math or language arts, typically an assignment is given and the students then apply themselves to the work. The teachers circulate to answer individual questions and look over student progress, rather than going over the parts of the assignment with the entire group and calling on individuals to recite.

Teachers' comments are carefully worded to be as encouraging as possible. Absolute negatives were never used if a student had attempted to provide a solution or answer a question. Clint and Brenda used phrases such as "Not quite"; "Think about it a little longer"; "Can anyone help clarify this"; "Good try"; "It's a tough one" when a student responded wrongly. Or, they simply said nothing and moved on to the next student. Since they consistently gave positive feedback for correct response, absence of confirmation served as sufficient indicator that the response was not correct.

During the period I was observing at the school the sixth grade was focussed on two major activities. In the first five weeks the class was preparing for its week-long trip to the beach, the culminating adventure that all Lower School children look forward to. Each student was preparing a mural, oral presentation, and written report on a topic relating to the natural or cultural history of the ocean region they would visit. The weeks following the trip were directed toward completion of the school year, and included finishing oral reading books, participating in a fifth-sixth grade musical production, completing shop projects, writing evaluations of their sixth grade experience, and taking the required achievement tests for seventh grade placement.

### A Day in the Sixth

The school day in the sixth begins at 8:30. The teachers are available early for individual help and some students arrive and work on due or overdue assignments. Others wander in during the first 10 or 15 minutes. If Clint has not provided a summing up at the end of the previous school day, then the announcements include an assessment of "how we're doing" in general. This call to order is conducted from the teacher's speaking chair (see Figure One), the locus of most announcements that the students are to pay particular heed to. During these first minutes Brenda sits at the teachers' desk receiving student work and checking it off in her ledgers. She adds homework and scheduling updates to the announcements, frequently pointing out the names of those delinquent in homework that she has written on the blackboard. Often Clint or Brenda leads the class in self-correcting, or having a neighbor correct, a math or language arts assignment. These are then also submitted to the "in" basket on the desk for her to look over and record. Throughout this period and at many other points in the day students freely leave their desks to ask Brenda quiet questions about the status of their work or to request help or clarification.

The first 50-minute period is devoted to language arts. Four days a week the entire class participates; one morning a week the class splits into two groups: advanced French students go to oral reading groups; low French students continue with language arts. The teachers find a strong correlation between slow learning in French and in the writing, grammar, and spelling required for language arts.

The language arts class may be an individual writing period or a group-focussed activity. Group work during the language arts period included going over homework sheets on points of grammar and punctuation, with discussion of the "whys" for each answer; dictation of a list of words associated with the upcoming trip to the ocean; and a team competition game, "spelling baseball". Composition work comprised most of the language arts sessions during the period of observation. Some writing assignments were large, ongoing composition tasks; others were short tasks to be completed that day. When the observation period began the students were working on a character sketch and sessions were devoted to drafting, revising, and cleaning up final drafts. Later students were given some language arts time to work on

their major reports for the ocean trip. Short assignments were in essay, letter, and brief answer form, including a follow-up letter to the state governor thanking him for meeting with the class; a short story; a book review; an essay comparing a novel and a short story that they had previously read; a thank you note to someone who had helped the student through the year; a list of books read, with a comment about each; the students' self-evaluations of their year's work; the list of ways to succeed for the incoming sixes; and oral performance tasks including recitation of memorized poem.

For composition work, the teachers set an assignment, outline its requirements, and then permit the students to proceed with it at their own pace, circulating among them to offer assistance, advice, and direction and allowing students to come to them for private consultation. Generally, this pattern of one-on-one work with students prevails over group recitation in all Clint and Brenda's teaching. The teachers also rely extensively upon modelling good student work to motivate others and to clarify the direction their classmates' work should take. Brenda recalled that, on their second major assignment of the year, several students came to her and Clint and asked permission to redo their work after they had heard two classmates read their essays aloud.

They just realized what we wanted and what they should be able to do. It's (listening to classmates' work) the best way for them to know what they can do.

The record of one such session midway through the development of the character sketch will illustrate Brenda and Clint's instructional method:

- o Having seen their drafts, Clint requests that two students read their essays to the group. Others pull out their papers and a third volunteers to read; a fourth is called on and performs. Each reading is followed by evaluative comments, generally positive, but often also directive; other students pay close attention to these remarks.
- o Following the first reading, other students begin to converge on Brenda's desk to receive their drafts from her or, with their own papers, to pose quiet, earnest questions about the character sketch assignment.
- o Following the fourth reading, Clint announces, "If you haven't finished your character sketch, do that now." Students begin writing at their desks, others move to Clint and Brenda for questions and help. Some, like the four readers, proceed to work on revisions of their full drafts. Clint has up to eight students waiting for private consultation at the teachers' speaking chair; each is handled in turn, in quiet conversation.
- o Brenda pauses in answering individual questions at her desk to clarify the assignment, asking the general group: "Do you know what the assignment is?" Some, having been absent on a group trip, do not. She goes to each of them.



- o Having whittled down his queue of questioners, Clint begins circling the room, pausing by most desks to speak, quickly or at length, with the writers. His circuit of both student circles takes over 10 minutes. He pauses several times to make statements to the whole group, clarifying the assignment, e.g., noting that the teachers are not correcting spelling, since it is just a draft, but that these concerns will come later.
- o Clusters of students re-form at Brenda's desk and, for Clint, where he is seated next to a writing student. They take questions in turn, then again circulate the room, making some evaluative comments loud enough for the whole class to hear. For example, Clint tells one student, "That's a pretty good physical description [of the selected character]; now what kind of personality are you going to give?" Some subsequent student questions relate to this same topic, modelled on their classmate's work.
- o The session ends with Brenda outlining for the whole class the next steps in the writing process and evaluating their work. Now that they have completed the physical description and characteristics they will need to outline a story about their character. This should be done in homeroom or recess periods or at home and is due to be shown to one of the teachers in the next two days. Work seen so far is rough; spelling, handwriting, punctuation are not good enough. Brenda recommends students give their drafts to a classmate for proofreading before submission to the teachers.

The next class period, from 9:20 to 10, is French and reading, with the class divided into two achievement groups on the basis of their proficiency in French. French I students are those who are new to Mount Stanton (and therefore to the French program) and those who have not achieved well in the language in past years. French II students are ongoing Mount Stanton students with higher proficiency. Placement is determined by the French teacher's test. The French class, held in the auxiliary classroom building, is, unlike Clint and Brenda's homeroom classes, highly group structured and performance-oriented. The French teacher queries students randomly in the class and they are expected to respond by reciting in French. The class is extremely unpopular with the sixes. Their motivation is low and they regularly fail to prepare their assignments. The French teacher writes a separate report for the students' mid-year and final evaluations.

The half of the class not attending French are divided into two small groups for reading, each led by one of the homeroom teachers. Reading is taught entirely orally. Students may select between the two books offered for their reading circles, so the numbers may be somewhat uneven in each group. One group convenes in the library and the second in a small circle of chairs within one of the classroom seating circles. Each member, including the teacher, reads a section in turn. The students making up of French II groups are generally quicker and more fluent in their English reading as well. Brenda and Clint sometimes select more challenging books for those groups. However, the French I groups include the 11 new students, many of whose

language arts ability is high. Thus, the achievement grouping is modified, so that as many as half the members of the French I reading groups are in fact high level readers--for Brenda and Clint a critical aspect of the reading program's success. Each group, in their view, must have some good readers to model for the others. For the low French group, "reading is the most important class of the day," Clint states.

For Clint and Brenda oral reading fulfills a variety of important educational goals. It offers them plenty of opportunity to assess the students' progress. The oral performance trains the students simultaneously in reading skills (scanning, vocabulary, grammar) and in self-presentation skills (articulation, presence, pace). Finally, it illustrates to the children that, in Brenda's words, "there are a variety of ways to be successful."

It brings out strengths in different kids. Gayle was a horrible reader. She thought that she couldn't read at all and she said in her (self-)evaluation, "I found out that everybody else had to learn how to read, too." So it's sort of a humbling thing for some and a learning thing for others. Plus the goal is so obvious. They know from the sound if somebody's good, so they know what to work for; the role model is fairly defined.

After a recess period from 10 to 10:20, instruction resumes, with the class again divided into smaller groups for different subjects. During the observation period, the sixes were divided into "A", "B", "C", and "D" groups for this period, determining, for any day, whether they attend science, art, or math class. Both Clint and Brenda maintain that the students are not ability- or achievement-grouped for these subjects. The current grouping had been established shortly before the observation period was begun, however, in order to provide extra math catch-up time for a subset of the students who had been away two weeks on a school-sponsored trip. Two groups jointly attend Clint's math class and the other two small groups go off to the classes taught by the subject matter specialists. Like the French teacher, these teachers write their own prose evaluations of the sixes for mid-year and final reports.

For math class, the 20 or so students take places in the seating circle next to the blackboard, where Clint usually presides. The math period often includes some presentation on the teacher's part, student desk work on problems (self-corrected or corrected by a classmate), questions from students, recitation, individual consultation between the teacher and students, and student-student cowork. If Clint chooses to go around the circle asking for answers to problems just set or assigned as homework, students may decline to answer without providing any excuse. If some class members appear lost, Clint frequently requests that their neighbors assist them with a private explanation, showing how they themselves arrived at their solution. This is the best method, Clint says, for they both learn from the experience, the slower student getting more help--specifically, help from someone who has just learned the process--and the quicker student forced to analyze how he/she solved the problem.

Between 11:30 and 1:10 the sixes eat lunch and take their physical education class. In order to provide separate classes for the boys and girls, Clint assists the gym teacher by taking half the class himself. He regards this as very important work. Growth in physical ability and willingness to compete and cooperate in physical activities are crucial aspects of the Mount Stanton education, he explains. Further, participation in PE enables him to monitor student progress. Like the oral reading performance task, PE provides an important, alternative avenue to success for certain students. PE performance is noted on the mid-year and final student evaluations written by Clint and Brenda.

The half-hour following PE is a "cool down" period. The entire class reconvenes for their most passive portion of the day. Clint and Barbara read aloud to the group or show films, or occasionally a speaker comes to address the class. This transition period is critical, they believe, so that students can settle down emotionally from the rigors of competition and exertion and gradually refocus their energy. During the first weeks of the observation this period was used to show films about the ocean and marine life and to read Ernest Hemingway's *THE OLD MAN AND THE SEA*, preparatory to the beach trip. Students are expected to take notes during the films and to answer questions about films and oral readings at their conclusion.

For the final two instructional periods of the day, from 1:40 to 2:20 and 2:20 to 3, the students are divided into equal thirds among music class, shop class, and homeroom period. Again the groupings (this time "reds", "greens", and "blues") are not related to achievement or ability. Brenda oversees the homeroom students who are working on major projects and on homework. This is a time for in-depth attention from the teacher, should students require it. Before the beach trip students worked on their murals, visited the library for background reading, and worked on their written reports. As the school year drew to a close, they put their time into tasks that needed to be finished up, many requesting extra time to go to the shop to complete their wood projects.

Clint spends the last two periods overseeing the work of the non-homeroom groups. He always drops in on the music class, "just to see how they're doing." The music and art teachers provide their own prose mid-year and final evaluations of the students. Most of Clint's time is spent at the shop, providing additional help to the shop teacher. The sixes have the option of working on a lawn mower motor-powered wooden go-kart, a project that Clint enthusiastically leads. Such projects, he believes, provide opportunities for students to really test out their abilities to organize, plan, and carry through a complex task. Shop, in general, he argues, is important as a vehicle for expressing different kinds of ability from that usually demonstrated in the classroom. Shop, too, is a separately evaluated subject, but, unlike all other subjects, it is assessed with a form, with only brief individual comments (See appendix). Students are rated on a four-point "strong" to "weak" scale for behavior, motivation, and achievement. Although shop is not conducted by the homeroom staff, if the students have elected to work on the go-kart project, it appears on Clint and Brenda's final evaluations.

If time permits, Clint provides an assessment of the day's activities after the class convenes for announcements and dismissal. The teachers require that every student's work area be cleaned up before anyone can leave. Many of the students stop to give the teachers a hug on their way out of the room.

Students whose work is overdue may be required to stay after 3, as "members" of Brenda's "After School Club". They are ordinarily given a day's notice, so that they can arrange their transportation (or, perhaps, rush to finish up). Notices for the After School Club and for those required to stay in and do homework during the morning recess period are written on the blackboard for the opening of school. No excuses are accepted once detention has been assigned.



## Assessment Records

### Permanent Assessment Records

Permanent school records at Mount Stanton Lower School include two general types of assessments, teachers' mid-year and year-end evaluations and, starting with the third grade, standardized achievement test scores. The teacher evaluations are sent home for parents and students to read and respond to. The test scores are available on request, but otherwise not distributed to parents or students.

Currently the school uses the SRA Achievement Series each spring for students in the third grade and beyond. The sixes completed language arts and math sections of the Level F/Form I version during the period of observation. Overall school scores on these tests are made available and are often of interest to parents considering enrolling their children at the school.

Twice yearly each teacher prepares an individual evaluation of the students. The first evaluation is sent home in January or February, at which time a parent conference is requested by the homeroom teachers. The second is completed just at the end of the school year barring serious problems such as possible retention of the student, year-end conferences are held only on parents' request. With the exception of the shop teacher, who uses a scaled form, each teacher prepares a prose report (See appendix). Mid-year and final evaluations are parallel in form and content.

For the sixes, the evaluation reports are made up of an essay jointly composed by Brenda and Clint; briefer prose evaluations from the teachers of science, art, music, and French; and the shop teacher's rating form. At mid-year Clint and Brenda's evaluation essays run about 900 words; at year-end they are approximately 750 words in length. These reports reflect the homeroom teachers' particular concerns, as well as their responsibilities.

In the final evaluations, the first paragraph outlines the student's performance during the week-long trip to the ocean, addressing such issues as interactional behavior and willingness to learn experientially. The content of this paragraph is developed jointly by the two teachers during their return trip from the ocean and draws on the evaluations of the tent counsellors who accompany the class, as well as the teachers' own observations. In some cases a second, short paragraph points out other behavioral issues which Brenda deems important.

A paragraph is devoted to each of the homeroom subjects. Brenda composes the sections on language arts and reading; Clint writes the math evaluation. He also comments, at some length, on physical education performance and, if the student has participated in the shop go-kart project, the success of that endeavor is briefly noted. A final paragraph sums up overall personal and academic development.

## Teachers' Assessment Records

The sixth grade teachers keep careful, complete, and current records on each student's work. Brenda maintains two file boxes of student work, sampled throughout the year. One box contains scored objective work, such as math problem sheets and grammar, punctuation, and spelling work sheets with teacher comments. The second consists of student writing samples. She uses these materials, especially the writing samples, to assess for herself the progress the class is making, as a group and as individuals. She also finds the material valuable in working with the students (and their parents) to show them how far they have come. At year end all the work is given to the students to take home—a record of their year's accomplishments.

Homework and classwork is consistently checked and scored and/or commented on. Brenda works every morning during the opening minutes of the class day to check in all incoming work, announcing who is overdue orally or through lists on the blackboard. Students come up frequently throughout the day to look at the check-off list, sometimes going off to rummage successfully in their desks or cubbie holes for the missing assignment. Brenda's check-off lists indicate only what is completed, not whether it was punctually submitted. Each paper must be dated, however, creating a record of the timeliness of students' work.

In the early weeks of the school year homework completion reports are sent home for the children's parents' signature and comments. These report-back sheets are reinstituted if the students lapse in their personal responsibility for their own work.

Assessments of major projects are recorded on a special form that Brenda has devised to reflect all components of the work (See appendix). These evaluation summaries are distributed to the students at the completion of the project and are also kept for Brenda's files until year end. They are a major resource for composing the evaluation reports. For the oceanography project, Brenda's summary of the classwork was supplemented with a teacher-authored evaluation of the students' actual behavior during the trip and a brief assessment of each student written by their trip counselor.

Performance in the oral reading circles is also recorded. Occasionally Brenda does an assessment of group members after the reading session, noting on a check-off form the general level of the students' achievement that day in skills such as vocabulary, sounding out words, assurance, flow, and comprehension. These records, too, are a valuable reference for preparation of the evaluations, as well as monitoring student progress.

## Range of Behaviors Assessed

### Making Citizens

Clint: About half our job is making citizens out of them...

Brenda: That's right.

Clint: ...and it isn't working if they're just citizens with us.

Brenda: Mhm-mhm. They have to go out into the wider world with it or what we're doing isn't really working....Like they will say, well, "I don't like her," and we say, "You don't have to like her, but you have to be respectful and you have to do all right with her."

In the Mount Stanton sixth grade, character development is equally as important as academic development. Clint and Brenda view the last year of Lower School as a transition point for their students. They are helping adolescents, beginning to address new issues in their own lives, starting to redefine themselves as separate, independent persons. And they will be entering a new educational environment. The Middle School teachers "never know the kids so well." This is their final year of genuinely intimate relations with a teacher. Thus, they believe it is imperative to model appropriate adult-adult behavior with their students, before they go off to more autonomous environments and relationships. Says Clint:

To leave out ethical and moral things at the very time kids should be learning these things is the biggest error. You might as well leave out reading, because it can't be more important.

When Brenda and Clint describe their goals for their students, it is invariably in terms of personal growth. During our first discussion of the classroom performance assessment study, Clint offered the assessment of the ocean trip as an example of how he and Brenda work:

I'm always assessing them. I look for the ability to get along, to do independent work, to take initiative, to solve problems.... (I try to) assess different skills, let different aspects emerge...through traditional activities and non-traditional ones as well. Building go-karts is a good example.

Clint and Brenda measure their teaching success as change in students' everyday behavior. Clint counts this past year a success, for they transformed a group whom the whole school considered "unruly, governable" into a supportive group of kids. He recalls that at least a half-dozen of the boys came up to congratulate one of their classmates on an excellent presentation: "It's such a healthy thing when kids can say

to each other, 'That's a nice job', or 'Boy, that's really neat.'" Brenda and Clint also value the sixes' ability to interact with comfortably with adults. They describe Ulrich, a somewhat troubled, and trouble-making, early adolescent:

- Clint: Ulrich's the kind of kid who doesn't have all the skills he needs yet, but he's trying to find out how the world goes together. But he'll be fine.
- Brenda: He's a good, affectionate, caring person and he wouldn't be in another school. He'd be raising hell.
- Clint: They'd squelch him.
- Brenda: For a kid like that to be able to come up and give you a hug at that age, it's pretty phenomenal, I think.
- Clint: He'd easily be on drugs....

The teachers argue that certain levels of interactional skill and personal confidence are prerequisite to successful functioning in their classroom. Thus, for some students, progress in these areas must precede academic growth. They describe Harold, who even at the end of the school year was noticeably less interactive than most in the class:

- Clint: Harold is a bright kid....He had pretty good academic skills, but he's not going to use them, do anything for you unless ...you get him chuckling. So having a good time about coming to school is about as far as we could have gone with that.guy.
- Brenda: To have him able to write as sincerely as he does and to have him think about his emotions and to have him think beyond himself as far as emotions are concerned, those are things that, well...He was the most unpleasant kid I ever met in my life....
- Clint: ...he's a different person. And it's not academic, like you're (researcher) talking about, but it does tie in....We went with him with math from when he didn't understand something, that he just would not listen. And he went from that point to "Would you show me how to do this? I really need to know how to do this now." That's 180 degrees.

They find that self-esteem and self-confidence are essential bases for academic learning. Clint offers Gayle, a young, very quiet class member as an example. In previous years he had taught her elder sisters, both of whom were



exceptionally bright. Gayle, however, came to the sixth with a history as a marginal student:

She always felt stupid. Now she feels better about her whole life. Success in reading has spilled over into her whole life....That's our goal. I don't care if she's the world's wizard at math or writing or anything else. Because she's come so far this year. And in the next couple of years, the next part of her life, she's going to feel so much better about herself that she's getting there quicker.

When they used to give pre- and post-testing in the sixth, Brenda recalls, it was these very students with whom they stressed non-academic growth who had often leapt furthest academically, sometimes as many as four grade levels.

#### Assessments of Personal and Social Characteristics

Clint and Brenda dedicate two to three paragraphs of their final evaluations to assessments of social and personal characteristics. Table One summarizes their comments from those sections of the reports. Personal skills were most frequently noted (319 comments, averaging more than 8 comments per student). This is consistent with the teachers' self-reported emphasis on personal development.

Over half of the remarks on personal skills address just five traits: willingness to take risks and try new things; general humor; sense of confidence or self-worth; overall kindness, thoughtfulness, and niceness; and level of ability and/or willingness to function as an independent person.

Risk-taking, confidence, and independence were assessed for every student. They are attributes that the teachers consider fundamental to success in their classroom, academically as well as personally. Following are typical assessment comments, illustrating how these three attributes intersect as the basis for learning:

Arthur had a good time on our beach trip. He was interested in everything, wanted to give most things a try...

We loved Bill's enthusiasm for the beach trip experience and appreciated his aggressive interest in all that was going on. He learned a lot and spent days with his eyes wide open, always receptive to opportunities.

Emily appreciated the independence and trust she received on our beach trip and responded reliably, taking care of others as well as herself...We would take Emily anywhere and hope she keeps us in mind when she reaches [beach trip] counselor age.

TABLE ONE

Final Evaluations: Assessments of Personal & Social Characteristics<sup>1</sup>

| <u>Personal Skills</u>                       |     | <u>Social Skills</u>                       |     | <u>Academically Related Skills</u> |     |
|----------------------------------------------|-----|--------------------------------------------|-----|------------------------------------|-----|
| Risk taker/tries new things/varied interests | 41  | Group membership <sup>5</sup>              | 35  | Responsible/self monitoring        | 62  |
| Good humor                                   | 39  | Appreciative of others                     | 21  | Task completion                    | 37  |
| Confidence/self-worth                        | 32  | Good sport                                 | 18  | Serious about work                 | 33  |
| Kind/nice                                    | 30  | Relates to adults                          | 15  | Volunteer to perform/assist        | 23  |
| Independent                                  | 20  | Affectionate                               | 15  | Punctuality                        | 19  |
| Bad habits                                   | 16  | Polite/courteous                           | 13  | Effort/attention                   | 16  |
| Coordinated/athletic                         | 15  | Respected by group                         | 12  | Organized                          | 9   |
| Attitude <sup>3</sup>                        | 12  | Relate to teachers/seek help appropriately | 10  | Enjoy learning                     | 8   |
| Uncooperative/insist on own way              | 11  | Gives of self                              | 8   | Study habits                       | 5   |
| Leads                                        | 11  | Makes friends                              | 6   | Knowledgeability                   | 5   |
| Independent thinker                          | 9   | Popular in group                           | 4   | Ask questions                      | 5   |
| Competitive                                  | 9   |                                            | 157 | Other <sup>6</sup>                 | 3   |
| Follows instruction or correction            | 9   |                                            |     |                                    | 225 |
| Mischievous                                  | 6   |                                            |     |                                    |     |
| Sincere                                      | 6   |                                            |     |                                    |     |
| Sensitive/deep feeling                       | 6   |                                            |     |                                    |     |
| Respect rules or boundaries                  | 6   |                                            |     |                                    |     |
| Balanced personality                         | 6   |                                            |     |                                    |     |
| Loyal                                        | 5   |                                            |     |                                    |     |
| Forthright/direct/clear                      | 5   |                                            |     |                                    |     |
| Good citizen                                 | 5   |                                            |     |                                    |     |
| Decisionmaker/accept decisions               | 5   |                                            |     |                                    |     |
| Emotional maturation                         | 4   |                                            |     |                                    |     |
| Other                                        | 11  |                                            |     |                                    |     |
|                                              | 319 |                                            |     |                                    |     |

<sup>1</sup> Stated positively or negatively, according to majority of remarks.<sup>2</sup> Use obscenities, rumor-monger, fighting, spitting, game playing.<sup>3</sup> Positive attitude = 8, negative = 4.<sup>4</sup> Judgmental, complaining, creative/imaginative, emotional.<sup>5</sup> Cooperative/group oriented = 20, tight with a few friends = 8, loner = 6, uncooperative/not group-oriented = 1.<sup>6</sup> Self-motivated, slow worker.

n = 38

Patti has done some changing, we think. She seems more able to handle problems on her own--social, physical, and mental. She is more confident in her abilities, more willing to try new things, more able to stick with tasks.

We have noticed many achievements with Gayle since September and these were most apparent on the beach trip. Her insecurities are at an all time low as confidence urges her into more risk-taking adventures and learning experiences....Often, she was independent and competent about responsibilities, chores, possessions, et al.

While humor and kindness relate closely to the social skills that Brenda and Clint assess, they are seen largely as personal attributes, determining how much the students actually get out of the experiences offered to them.

Kim ...enjoyed the (beach) trip while maintaining her consistent, mature behavior. Kim took Ilone under her wing....Kim also had the intuition to know when help was needed in the kitchen. She was a reliable worker, never complained, and cheerfully accepted any task put before her...We hope we can convince her to make the trip again in a few years as a counselor.

Throughout the (beach trip) week Ali was good-humored, loving, positive, and tough. She wanted to try it all and learn as much as possible.

A wide range of specifically social skills (157 comments) were also commented on in the final evaluations. Almost all evaluations (35 of 38) contained an assessment of the student's attitude toward group membership and cooperation with other students and with adults. In addition, students' ability to interact with adults--the teachers and adults not well known to them--was also assessed in 15 of the 38 cases. Appreciation (21 comments), affection (15), and politeness (13) all appear as important aspects of social development.

Among academic task-related skills, taking responsibility for oneself again appears as a critical attribute, in this case accounting for over 25 percent of the evaluative remarks (62 of 225 comments). The closely affiliated characteristics of task completion and seriousness about work account for another 30 percent (70) of the comments. Together with punctuality (19 comments) and level of effort (16), attention and completion characteristics account for almost half (105) of the academically-related assessments. Volunteering in class and in other work settings is notably frequently assessed (23 comments), consistent with the importance placed on cooperative and group-oriented behavior generally.

### Assessments of Reading

Tables Two and Three display assessments of reading, in the final evaluation and in a sample reading circle session, respectively. Oral reading performance skills are most frequently assessed in the final evaluations. Table Two shows 114 comments on performance, 60 on comprehension, and 51 on personal and social skills. Among social and personal characteristics, level of attention or effort is considered most frequently (21 comments). Appreciation of literature and the ability to learn from literature is also heavily weighted (17 comments). Assessments of comprehension are almost evenly divided between general comprehension skills (27 comments) and knowledge of vocabulary (30). A variety of oral reading skills are assessed. Smoothness and phrasing (38 comments), scanning (9), and use of punctuation (9) serve to define the students' ability to interpret grammatical structure as they read. Thus about half of the evaluations of oral performance report progress in understanding of syntax. Assessments of pacing (20 comments), i.e., regular and unlabored pronunciation of words, and sounding out words (8) provide evaluations of vocabulary recognition and production. These constitute about 25 percent of the oral performance assessments. In addition to lexical and grammatical skills, significant importance is attached to oratorical skill: expression is the second most often assessed performance skill (21 comments).

In every evaluation, areas of progress and areas needing improvement are balanced. For example, one of the slower readers is assessed as follows:

Mark is still working on punctuation awareness and accuracy, but has made much progress since Fall. He maintains a better pace, sound words out more easily, and is a more accurate reader. Mark still needs help understanding difficult vocabulary and benefits from any reading possible....any effort in reading over the summer months would be time well spent.

A greatly improved student is described thusly:

Steve is a good reader now; accurate, appropriately paced, and expressive. It is much easier to understand him, so the emphasis on enunciation and the regular practice in reading aloud have paid off. He seldom omits small words and pays closer attention in class, bringing forth improved comprehension.

For an accomplished reader, higher level skills are assessed:

Tim woke up to the pleasures of literature and developed likes and dislikes while evaluating books seriously. He appreciated the skills he learned in reading aloud and has new respect for this task that is not as easy as it appears. Tim is an accurate reader, smooth and well paced. He possesses a vocabulary beyond his years and is always interested in learning new words.



TABLE TWO

Final Evaluations: Assessments of Reading

| <u>Performance Skills</u>    |          | <u>Comprehension Skills</u> |          | <u>Personal &amp; Social Skills</u> |          |
|------------------------------|----------|-----------------------------|----------|-------------------------------------|----------|
| Smoothness/phrasing          | 38       | Vocabulary                  | 30       | Effort/attention                    | 21       |
| Expression                   | 21       | Comprehension <sup>2</sup>  | 27       | Appreciate                          |          |
| Pacing                       | 20       | Other <sup>2</sup>          | <u>3</u> | literature                          | 17       |
| Scanning                     | 9        |                             | 60       | Respect for task <sup>3</sup>       | 6        |
| Use of punctuation           | 9        |                             |          | Other <sup>3</sup>                  | <u>7</u> |
| Sound <sup>1</sup> out words | 8        |                             |          |                                     | 51       |
| Other <sup>1</sup>           | <u>9</u> |                             |          |                                     |          |
|                              | 114      |                             |          |                                     |          |

n = 38

<sup>1</sup>Enunciation, dialect reading, omit words, breathing irregular.<sup>2</sup>Memory for detail, use of contextual cues.<sup>3</sup>Anxious/defensive, share ideas, pride in work.

Table Three offers a protocol of all the verbal comments and corrections taking place during one 40-minute reading session. The students were well along in novel about the Civil War, a text that was challenging in vocabulary, in use of dialect, and in story line complexity. Oral performance evoked the most teacher comments (59), both correction of errors and requests for improved performance. Comprehension comments were second most frequent (23 comments). Personal and social behaviors received 15 comments from the teacher, including four calls for attention from the group or from individuals, three comments of encouragement during a student's turn at reading, and eight post-reading positive assessments on a reader's effort. In addition, on five occasions other students in the group made encouraging comments to a struggling classmate.

By far the most frequent correction was for mispronounced words (42 comments). Corrections of stress and rhythm (7) reflect failures to comprehend grammatical structure, including punctuation markers. Students were not permitted to read without emphasis and expression; this was corrected 5 times. In some instances this correction also appeared to be a check on or correction for interpretation of syntactic structure. The teacher did not always follow mispronunciations of words with comprehension checks. She requested synonyms or definitions of words only 13 times during the entire session (compared to 42 pronunciation corrections). The teacher checked general comprehension 8 times, asking for explanation of the paragraph or sentence just read 6 times and twice evoking group discussion of the context for the section being read. In addition, she twice provided background information on events or personages of the Civil War era.

In both the final evaluation report (Table Two) and the sample reading session (Table Three), performance variables are the most frequent focus of teacher assessment. Corrections of rhythm and stress and of expression appear in the class session protocol; in the final evaluations these tests of syntactic comprehension are reflected in assessments of phrasing, expression, and scanning ability, as well as overall comprehension skill. In class and in the report vocabulary is afforded the most frequent comment: During class, 42 pronunciations are corrected and 13 words' comprehension is checked. The evaluation reports contain 30 comments on vocabulary comprehension ability and an additional 8 comments on ability to sound out new words.

TABLE THREE

Verbal Comments & Corrections: Reading Circle Session<sup>1</sup>

| <u>Performance Skills</u>             |    | <u>Comprehension Skills</u>                     |    | <u>Personal &amp; Social Skills</u>                       |    |
|---------------------------------------|----|-------------------------------------------------|----|-----------------------------------------------------------|----|
| Word mispronounced                    | 42 | Give synonym                                    | 7  | Positive comment on performance                           | 6  |
| Correct stress/rhythm                 | 7  | Explain word                                    | 6  | Attention requested                                       | 4  |
| Add enthusiasm/expression             | 5  | Mispronounced correctly                         | 3  | Group                                                     | 2  |
| Word omitted                          | 2  | pronounced                                      | 3  | Individual                                                | 2  |
| Add dialect                           | 2  | Explain paragraphs                              | 4  | Encouragement during reading                              | 3  |
| Reduce dialect for clarity of meaning | 1  | Restate sentence in other words                 | 2  | Positive comment on attitude toward correction/repetition | 2  |
|                                       | 59 | Relate to past parts of book (group discussion) | 2  |                                                           | 15 |
|                                       |    | Provide background information on topic         | 2  | Encouragement from classmates                             | 5  |
|                                       |    |                                                 | 23 |                                                           | 20 |

<sup>1</sup> Protocol taken of a 40-minute reading period devoted to Across Five Aprils, a Civil War saga among the most challenging books used in the sixth grade. Students are from the French II group.

### Assessments of Language Arts

Language arts final evaluations are focussed primarily on writing, but some attention is given to oral performance for this subject as well. Table Four summarizes the teachers' assessments on the language arts section of the final evaluations. Technical writing skills are most frequently assessed (144 comments), although higher order writing skills are also given extensive review (106 comments). Only 25 comments about oral presentation skills are given. Personal and social skills are afforded 66 comments and, as in the report sections on reading-related behaviors, effort is given the most attention (24 comments, over a third of all remarks). Effort is also reflected in assessments of attitude toward reading (14 comments) and, in part, in the 14 assessments of pride in work. The teachers' emphasis on interactional skills is reflected in the 9 comments on students' help-seeking.

Although only three class sessions were devoted to spelling instruction throughout the observation period, spelling is the most frequently noted technical skill, receiving almost one quarter (34) of the 144 comments. Proofreading, a practice highly emphasized in the teachers' classroom interaction with the students, is the technical skill second most often assessed (29 comments). In class, proofreading and pride in work were closely associated by the teachers. Penmanship is commented upon 22 times, also an accurate reflection of classroom emphasis. Assessments of the mastery of grammar (22 comments) range from accurate capitalization to understanding of syntax and derive, at least in part, from work sheet assignment performance.

Assessments of higher order writing skills stress organization (19 comments) and cohesion (17), together accounting for a third of content and style comments. Depth (14 comments), self-expression (13), and sincerity (7) in writing are valued more highly than knowledgability (4), suggesting that learning to write well and to enjoy writing are more important to the sixth grade teachers than communication of subject matter through writing.

Writing assessments, like reading assessments, balance progress made and progress needed, as well as expressing Brenda and Clint's objectives for student self-reliance and and self-monitoring:

It has been fun to watch Fran's progress in writing. Mid-way through the year, she developed more depth and sensitivity while increasing her effort on written assignments. She is more competent with organizing her thoughts and, with the exception of a few misspellings peppered throughout, reflects pride in all she does. The next step to maturity is proofreading. Many of Fran's errors are obvious ones that she could catch herself, so this should be a goal for next year.



TABLE FOUR

Final Evaluations: Assessments of Language Arts

| <u>Technical Skills:</u>           |     | <u>High Order Skills:</u>    |     |
|------------------------------------|-----|------------------------------|-----|
| <u>Graphic, Lexical, Syntactic</u> |     | <u>Content, Style</u>        |     |
| Spelling                           | 34  | Thoughts organized           | 19  |
| Proofreading                       | 29  | Cohesion                     | 17  |
| Penmanship                         | 22  | Depth/interest/              |     |
| Grammar <sup>1</sup>               | 22  | involvement                  | 14  |
| Punctuation                        | 15  | Writes for                   |     |
| Paragraphing                       | 6   | self-expression <sup>3</sup> | 13  |
| Dictionary use                     | 5   | Research skills              | 11  |
| Accuracy/precision                 | 5   | Directness/clarity           | 10  |
| Other <sup>2</sup>                 | 6   | Sincerity                    | 7   |
|                                    | 144 | Knowledgeability             | 4   |
|                                    |     | Creative wording/            |     |
|                                    |     | phrasing <sup>4</sup>        | 4   |
|                                    |     | Other <sup>5</sup>           | 7   |
|                                    |     |                              | 106 |

| <u>Verbal Skills</u> |    | <u>Personal &amp; Social Skills</u> |    |
|----------------------|----|-------------------------------------|----|
| Oral presentation    |    | Effort                              | 24 |
| ability              | 12 | Pride in work                       | 14 |
| Listening skills     | 4  | Attitude                            | 14 |
| Other <sup>6</sup>   | 9  | Appropriately                       |    |
|                      | 25 | seek <sup>6</sup> help              | 9  |
|                      |    | Other <sup>6</sup>                  | 5  |
|                      |    |                                     | 66 |

n = 38

<sup>1</sup> Apparently refers to both understanding syntactic structure and to specific markers for written syntax, e.g., punctuation, capitalization.

<sup>2</sup> Capitalization, sentence structure, integration of art work.

<sup>3</sup> Research skills are mentioned generically and specifically, including outlining, note-taking, paraphrasing, condensing and organizing material.

<sup>4</sup> Writes to report research, humor, fluency.

<sup>5</sup> Comprehension of verbally presented material, positive attitude to oral presentation, phrasing, verbal expression, oral vocabulary.

<sup>6</sup> Learn from homework, meet deadlines, follow directions.

As Ulrich matured, writing and self-expression became easier....His style is clear, direct and often peppered with humorous wit. We enjoy reading what Ulrich composes and respect the obvious growths he has made this year. He shows better organization of thought, more cohesion, neater penmanship and a broader understanding of basic grammar requirements. Ulrich is not the best of spellers but is quick to seek help and can even be cajoled into using a dictionary. His exuberance doesn't include proofreading but when he slows down a bit, he has the potential for catching many of his own errors....He feels rightfully proud of his obvious progress.

The teachers' integrated model of learning is also reflected in the evaluations. Reading and writing, oral presentation and written composition are tightly integrated skills, dependent upon one another. And all are tied to personal and social skill development. Their evaluation of Jim's progress in reading and writing illustrates their instructional philosophy:

Reading has new importance to Jim since he has discovered that reading aloud is not as easy as it once appeared. He has concentrated and worked hard over the year to read accurately, smoothly and with more expression. He is just transcending from smoothing words individually to phrasing and scanning ahead, which is part of the process. He still doesn't catch his own errors and needs more focus on comprehension, but he's getting better....Perhaps because Jim is just beginning to enjoy literature, his writing also reflects a borderline discovery approach of the same nature. He doesn't like to write and will be the first to admit it. He is bothered by spelling problems which, until lately, have inhibited his written expression. It has taken Jim a long time to build up enough confidence to write freely. He displays deeper thought, more cohesion and better organization than we saw earlier. His understanding of basic grammar is improving and he is able to stay focused on a writing assignment, which is fairly recent....He profited from the small research classes and used his time productively. Again, focus and maturity come into play.

Twenty students' oral presentations of oceanography project reports were observed during the study. Table Five reports the teachers' comments during and after these presentations. Together the written and oral versions of these reports constituted the largest language arts project during the observation period. Minimally, the teacher(s) responded to every report with a positive comment or a "thanks", if the report was brief or perfunctory. Positive comments included "That's the best report on (topic) we've had in years", "Very fine work", "An outstanding job", "Good work", "One of your best efforts to date", "Now we see how it's supposed to be done".

TABLE FIVE

Verbal Comments & Corrections: Oral Reports

| <u>Concluding Comments</u> |    | <u>Presentation Style</u> |    | <u>Content</u>    |    |
|----------------------------|----|---------------------------|----|-------------------|----|
| Positive comment           | 11 | Speak too softly          | 5  | Knowledgeable on  |    |
| Thanks                     | 9  | Make distracting          |    | topic             | 11 |
|                            | 20 | gestures/noise            | 3  | Conclusion        |    |
|                            |    | Use mural                 |    | follows from      |    |
|                            |    | appropriately             | 2  | report            | 4  |
|                            |    | Engage audience           | 1  | Able to answer    |    |
|                            |    | Bad posture               | 1  | questions         | 3  |
|                            |    |                           | 12 | Good explanation  | 2  |
|                            |    |                           |    | Data good, but    |    |
|                            |    |                           |    | low               |    |
|                            |    |                           |    | comprehension     | 1  |
|                            |    |                           |    | Comprehend        |    |
|                            |    |                           |    | hard topic        | 1  |
|                            |    |                           |    | Report to point   |    |
|                            |    |                           |    | of topic          | 1  |
|                            |    |                           |    | Expertise will be |    |
|                            |    |                           |    | applied on        |    |
|                            |    |                           |    | beach trip        | 1  |
|                            |    |                           |    |                   | 13 |

| <u>Provide Assistance</u> |   | <u>Directed to Audience</u> |    |
|---------------------------|---|-----------------------------|----|
| Ask directive             |   | Request attention           | 9  |
| questions                 | 3 | Individual                  | 6  |
| Help control              |   | Group                       | 3  |
| students' questions       | 2 | Good question               | 3  |
|                           | 5 |                             | 12 |

n = 20

For five students the teachers stepped in to provide assistance, either directing questions to the speaker, in order to get the talk back on track, or to help with the management of questions from the audience, when these interrupted the flow of the presentation. Comments to the audience were mostly requests for attention (9), but some students were also complemented for insightful questions (3).

In addition, the teachers commented on presentation style (12 comments) and content (13). Five students were told to speak up (during and/or after their talk). Three were admonished for making distracting noises or gestures while speaking. Use of the required visual aid was praised twice.

Content remarks were generally directed to the students' demonstration of mastery of the topic. Four comments evaluated the extent to which the conclusion of the report followed from the data or explanation presented. This corresponds to the teachers' frequent classroom queries of how students came to a conclusion, or how they know what they claim to know. For example, a film during the oceanography project focussed on Darwin's theory and the steps to his scientific analysis had been carefully discussed. In one-on-one work significant attention was given to logical progression of thought and creating an evidence base for claims.

Although group question-answer sessions were relatively infrequent, they did sometimes occur at the conclusions of films of story-readings in the post-PE listening session. In part, these questions clearly served to monitor student attention and to spotlight failure to be attentive. These were the only occasions during which Brenda and Clint chose to call on students who were least likely, rather than most likely, to know the answers. Although the post-listening sessions included higher-order questions, one typical series requested that six students recall facts from the story, three add to the story by developing background information, two restate the plot, and one explain why something had been done. That is, six questions were recall queries, two tested comprehension, three required interpretation, and one demanded analysis.

Brenda and Clint value the higher order skills highly and stress them in their individual work with students. However, they also believe it is important to expose the students to all forms and levels of questioning. During the year the sixth uses only one textbook, spending a few weeks working through a biographical section of Scott Foresman's IMAGE 1, including the questions for each story and the unit review. Analysis of these questions found 26 comprehension questions (38%), 17 application questions (24%), 14 recall questions (20%), 10 evaluation questions (10%), and three analysis and three synthesis questions (4% each). Unit review questions were restricted to recall and comprehension. In explaining why the textbook was used, Brenda stated that it is important that the students do even the true-false questions, since "they should be exposed to that; there's nothing wrong with it." Such forms of questioning are not regularly practiced in the sixth and, especially for students who are leaving Mount Stanton, rehearsal of these skills is of value.



### Assessments of Math

In contrast to the final evaluation assessments of reading and language arts, in which academic skills comprise the large majority of comments, in the math section of the evaluations personal and social skills constitute fully 40 percent of all comments. This directly reflects Clint and Brenda's belief that attitudinal problems, not lack of ability, are the source of most math failures. Table Six details comments on math achievement for 37 final evaluation reports. Personal and social skills are assessed with 112 comments, compared with 169 comments on academic math skills, 110 remarks, (40% total) on technical skills and 59 (20%) on higher order skills.

Parallelling the reading and language arts evaluations, in the math evaluations effort is the most frequently assessed personal characteristic (20 comments). And, as in those other subjects, attitude is also commonly noted (18 comments). However, confidence appears here for the first time as an academically-related attribute, with 19 comments, nearly as often assessed as effort and oftener than attitude. (In the final evaluation sections on personal and social characteristics confidence is the third most frequently assessed personal skill.) Performance under time pressure (14 comments), also a confidence indicator, is uniquely noted for math. The comment "interested in higher math skills" (10 comments) echos the reading assessments' comment "appreciate literature", both suggesting that learning is made easier and more fruitful by genuine interest in the subject. The math evaluations consider frustration with the subject as a significant factor (6 comments) and pride in work and appropriate help-seeking (6 comments each) appear here, as they do on the language arts evaluation, however less heavily weighted in math. The math evaluations also make mention of students' ability to ask good questions, closely allied to the ability and/or willingness to seek help.

Technical skills and higher order skills are carefully distinguished in the math evaluations. For example, one of the slower students is described as follows:

Math is still a weak area for Elizabeth, but she has learned new skills and sharpened many old ones over the year. She continues to be slack with multiplication facts and should work on building up speed with these. Her conceptual knowledge is pretty solid, but she is forgetful of processes or makes simple computational errors...

Genuine interest in mathematics is enthusiastically noted as explanation for academic improvement:

Ivan became more conceptual every day, not satisfied with finding the answer but interested in how processes were related. This is the thinking of a mathematician, and as his confidence soared, so did his test scores. It didn't phase Ivan whether or not he was under a time pressure as he set his own pace, used his increased speed in computations and appeared to be unconcerned. He was prone to careless errors when he hurried, like all of us, but did fine once he discovered he could work within the time frames.

TABLE SIX

Final Evaluations: Assessments of Math

| <u>Technical Skills</u> |     | <u>Higher Order Skills</u> |    | <u>Personal &amp; Social Skills</u> |     |
|-------------------------|-----|----------------------------|----|-------------------------------------|-----|
| Accuracy/carefulness    | 26  | Understand concepts        | 27 | Effort/attention                    | 20  |
| Mastery of content      | 23  | Understand process/        |    | Confidence                          | 19  |
| Testing performance     | 23  | can work in steps          | 24 | Attitude                            | 18  |
| Computational skills    | 15  | Problem solving            |    | Behavior under                      |     |
| Computational facts     | 8   | skills, esp. story         |    | time pressure                       | 14  |
| Speed                   | 7   | problems                   | 8  | Interested in                       |     |
| Proofreading            | 5   |                            | 59 | higher math                         |     |
| Neatness                | 3   |                            |    | skills (algebra,                    |     |
|                         | 110 |                            |    | geometry)                           | 10  |
|                         |     |                            |    | Easily frustrated                   | 6   |
|                         |     |                            |    | Appropriately                       |     |
|                         |     |                            |    | seek help                           | 6   |
|                         |     |                            |    | Competitive/pride                   |     |
|                         |     |                            |    | in work                             | 6   |
|                         |     |                            |    | Asks questions                      | 5   |
|                         |     |                            |    | Other <sup>2</sup>                  | 8   |
|                         |     |                            |    |                                     | 112 |

n = 37

<sup>1</sup>In three cases, problem solving ability was directly related to the student's reading ability.

<sup>2</sup>Appreciate relevance of math to other skills, under-value own math skills, complete homework, meet homework deadlines.

On the other hand, a negative attitude toward math, not lack of aptitude, is assessed as the primary cause of failure to learn:

Math is not as easy for Ginger as it would have been had she paid more attention in class. She started off the year with sincere effort but became easily distracted in class and missed much. She has the aptitude to be a pretty good mathematician if she set her mind to task. Many processes are just below the surface and come forward with simple refreshers, but others have gaping holes because she wasn't paying attention and didn't concern herself with the learning. Her lack of progress has been an attitude problem and one she needs to correct to maintain an appropriate standing in the seventh grade.

Among the math technical skills, accuracy and carefulness (26 comments), mastery of basic content (23), and performance in test situations (23) are most frequently noted. Remarks on mastery of computational facts (8 comments), fundamental computational skills (15), and notes on speed (7) and proofreading (5) make up the remainder of the technical skills assessments.

Higher order skills are assessed as "understanding concepts" (27 comments) and "understanding math processes/working in steps" (24). For eight students story problem-solving ability is commented on, and in three cases it is directly correlated to reading ability.

Testing behavior is not mentioned anywhere else in the final evaluations, although reading and language arts were time-tested as well as math. In addition, the math evaluation notes the likely math group placement for the seventh grade. Although final decisions from the seventh grade teachers had not been made at the time of writing the evaluations, the sixth grade teachers judge from the achievement test scores and their own knowledge of student performance (and their influence on the placement process) which of four math sections the students are most likely to find themselves in. (This topic is taken up again in the section on criterion-referenced assessment, below.)

#### A Developmental Base for Academic Success

Throughout their formal and informal assessments Brenda and Clint articulate their strong conviction that personal and social development are the prerequisite for academic success. As the excerpts cited above illustrate, they consider both achievement and aptitude in their evaluations. However, they tend to draw their basic distinction not between these two aspects of learning, but between achievement and aptitude, on the one hand, and personal readiness to learn, on the other.

Queried about the distinction between achievement and aptitude as they shape instruction in the sixth, Clint and Brenda acknowledged the distinction, but refocused the question on personal development. As is their habit, they

use student examples as the basis for their explanation of their pedagogy. Here Clint describes Steve, a bright, but immature and underachieving boy, and Linda, a mature and apparently highly able girl with a subtle learning disability:

Look at Steve and you know that he's intelligent and he has difficulties even writing a sentence. In that case you can't expect him to live up to his ability because he hasn't even touched it. And with him you're looking for progress...and when you see him reach it and then you feel good and you figure out what the next one is going to be and then you go for that one. But there are other kids that really go beyond their ability. Linda would be an example....

Steve remembers and he makes sense. He's thinking all the time. That's one of the difficulties with him, he's not paying attention in class because he's thinking on his own....He's got ability he hasn't used yet, but he's more able to use it, I think, because his communication skills, his feelings about himself are better. He's just a more confident person than he was, a more attractive person...

Clint and Brenda caution that personal and social skills, while they are fundamental to the learning that takes place in their classroom, are not reliable indicators of ability or achievement. Linda, mentioned in the passage above as working beyond her ability, is an example:

Brenda: Linda came to us with absolutely raving reviews because she's such a wonderful person. I never heard a negative comment about her ever. And it took a while to find out that the kid had some serious academic problems. Real, real holes.

Clint: Not just things she that she didn't know, but things that she couldn't learn in that way.

Brenda: And that's when, and then when you see that it's causing a child anxiety, "How come they got that and I didn't get it?", then it's time. That's when I call in (the school testing specialist) Dora and say, "What's going on here?" and Dora will run extensive, professional testing.

Consistent with this philosophy, the teachers' final evaluation reports stress personal and social development above all other assessments. Table Seven brings together comments on social and personal characteristics related to academic achievement from all sections of the final evaluations. A total of 507 comments on academically-related behaviors are made, an average of over 13 specific remarks per student. The six of the seven most frequently cited behaviors—effort/attention, responsibility/self-monitoring, confidence, attitude, seriousness about work, enjoyment of learning—are not specific to task achievement, but rather reflect orientation to learning and readiness to learn.



TABLE SEVEN

Final Evaluations: Assessments of Academically-Related Behaviors<sup>1</sup>

|                             |            |
|-----------------------------|------------|
| Effort/attention            | 81         |
| Responsible/self-monitoring | 62         |
| Confidence                  | 56         |
| Task completion             | 47         |
| Attitude <sup>2</sup>       | 46         |
| Serious about work          | 43         |
| Enjoy learning              | 35         |
| Appropriately seek help     | 25         |
| Punctuality                 | 23         |
| Volunteer to perform/assist | 23         |
| Pride in work               | 22         |
| Organization/study habits   | 14         |
| Follow directions           | 10         |
| Ask questions               | 10         |
| Easily frustrated           | 6          |
| Other <sup>3</sup>          | 4          |
|                             | <u>507</u> |

n = 38

<sup>1</sup>Sums assessments from all sections of the evaluation.

<sup>2</sup>Positive attitude = 31, negative = 15.

<sup>3</sup>Share ideas, self-motivated.

### Criterion-Referenced Assessment

Throughout their class work and their evaluation reporting, Brenda and Clint articulate their commitment to criterion-referenced assessment. During the two-months' observation period, the teachers were never heard making comparative evaluations of student performance or student work. They strictly maintained an attitude of evaluating each student on his/her personal progress, socially and academically. While they relied extensively on modelling among the students, they avoided situations in which student success was publically contrasted. For example, the oceanography topics were carefully selected so that better-achieving students received more challenging topics, thus enabling even weaker students to do a creditable job. Then some of the outstanding language arts students were asked to do the first presentations, modelling for the others what could and should be done. The oral reading circles also served to share performance and improvement publically.

However, student scores on math or language arts homework were never announced orally, even when papers had been exchanged for correcting. Each student simply dropped her/his assignment in Brenda's "in" box and could discuss the result privately. Ridicule of poor performance by classmates was never tolerated and rarely occurred. Compliments to others on a good performance were highly encouraged and occurred with surprising regularity.

One telling incident illustrates the values of mutual support and appreciation that Clint and Brenda encouraged in their classroom. Clint had prepared a 100-question multiplication exercise for a student who had requested additional help and practice. When the students arrived in the morning, some had seen it on the desk and requested that they be assigned a timed exercise of multiplication questions as a game and assessment of their own progress. Clint and Brenda agreed. Many of the students, especially those who arrived a bit late and had not seen the development of the exercise, took the "test" very seriously and some were very distressed when they failed to complete the 100 questions in the allotted time. A few minutes after the test had been corrected and many students were clustered around Brenda's desk telling their scores, three girls brought a fourth, tearful classmate through the crowd, dragging her by the hand. "Brenda", they said, "tell Patti she can do things. She's upset because she couldn't finish the test, but she can do lots of things!" Brenda immediately dropped all other concerns to attend to the distressed student, taking her aside privately to assure her of her skills, achievements, and math abilities. Later, Clint and Brenda spoke to each of the supportive students, complimenting them on their compassion and the way they had handled the situation. Each of these students' final evaluations notes their kindness and helpfulness.

Except for mathematics, the final evaluations contain few norm-referenced assessments. Of the over 700 comments on personal and social characteristics, only one implies relative standing in the group: One student is described as "as ready as any anyone we ever had" to succeed in the Middle School.

In the 38 reading assessments, there are 14 comments that suggest the general performance level of the student (5 "very good"/"fine"; 5 "good"/"able"; 3 "pretty good"; 1 "challenged"), but none that directly compares a student to the class. Similarly, in the language arts section, only three explicitly norm-referencing comments appear in the 38 evaluations. One student is described as having a vocabulary "beyond his years"; two students have spelling skills "above the sixth grade level". An additional 23 comments suggest achievement level: two students find vocabulary "tough"/"difficult"; one is a "pretty good" speller and two are "not very good"/"flawed"; eight students' writing is "natural"/"respected by others"/"notable"/"easy"/"good" and two students' is "slow"/"difficult". Reading and language arts assessments are almost exclusively couched in terms of personal growth since the fall.

In math, however, 29 of the 37 evaluations (71%) include explicitly norm-referenced comments. For almost all students, performance is generally characterized to suggest their place in the group. For 14 students, seventh grade math group placement is projected. For an additional 15 group standing is noted. For example, "Chester consistently scores above average on tests in class and on the standardized exam"; "Carl...is proficient at the class level"; David tests high in math, both in class and on the standardized exam; "To make progress which now places her in the higher average side of the class was no easy task (for Patti)." General characterizations of math skill level include "whiz"/"a natural"/"easy"/"high level" for five students; "capable"/"pretty solid"/"good" for four students; "challenged"/"difficult"/"hesitant" for eight students; "weak"/"area of concern" for six students, as well as 23 comments on specific skills that imply group-normed ability.

## Assessment Strategies

### Testing

Because Mount Stanton Lower School is a non-grading institution, achievement testing plays a minor role in the sixth grade classroom. It is so rare, in fact, that a testing situation evokes considerable anxiety among some students. As noted above, during the standardized achievement test two students broke down and were unable to finish and one student broke into tears over a pseudo-test of multiplication proficiency. Brenda and Clint work to avoid subjecting their students to test-like performances, arguing that the test situation does not offer a true measure of achievement or ability. They prefer to rely on assessments drawn from the normal progress of instruction. However, the teachers do choose to make use of self-designed tests for diagnosis and for sizing up.

The teachers make ready use of diagnostic testing when they suspect that a student may have an ability impairment. They called in the testing specialist for more than one student this past year, including a girl whose learning disability had gone undiagnosed throughout her Lower School years. This testing is done in private and is followed by meetings with the parents and the student and, often, parents and student together. A supplementary or auxiliary instructional program is jointly developed.

Clint composes a math test in the fall that the sixth uses for sizing up the students' achievement and ability in that subject. This test covers material they expect the average student would have mastered in the fifth grade and also skills that they anticipate having to teach in sixth. They refer to use their own test rather than the results of the fifth grade standardized achievement test because, as Clint puts it:

I want them tested right, tested with some kind of humanistic feel to it....I like it better to look at our own (test results), because I know how we set them up for those things. I just feel better about it if we've done it ourselves. I know that somebody like Ali [a dislexic], for instance, was unsettled....I would never want to get an impression of Ali from a test I hadn't given.

The problem with testing, Brenda adds, is that:

What you're getting in that (ordinary testing) situation is basically reading the directions and administering the test without any prefaces about what they (the results) are going to be used for...simply following the words in the test. Clint, when he does it, he really talks to the kids before about the test, not to put too much weight on it, that it's sort of a good thing to know how to do, but your life doesn't depend on it.



The sixth grade conducts required annual year-end achievement testing, this year administering language arts and math sections of a standardized exam on three different mornings. The test was announced and discussed several days beforehand. Clint and Brenda emphasized to the students that the test would be only one of the factors that affected seventh grade placement and that they, the classroom teachers, would have final approval of all placements and discrepancies from classroom performance assessments would be corrected. The standardized test, they repeatedly stated, was very unimportant compared with the students' ongoing classroom work. These assurances were repeated, and expanded, at the beginning of every test session. Clint went over examples and instructions carefully while making humorous comments to try to lower the tension. He also related how well the sixes always do on the test, far above the average, and how much of the material would be familiar to them already.

Clint and Brenda's pedagogy relies upon a relation of trust between teacher and student, in which the students learn to assess for themselves when they need help and grow to feel comfortable disclosing their need for help, to classmates as well as to the teachers. Clint describes his objection to testing this way:

Now if we were grading these kids, and their 75 or 90 [percent] were up on the door, that whole thing (trust) would be broken down. Then nobody—you know these kids at 12 or 11, they won't admit that they need help or that they couldn't do something as well as somebody else. But the whole thing is, well, this is life, and this is the kind of thing we're getting through. It takes down some barriers, it makes it so that people can learn what they need to learn and sort of feel that this is what life's about at that period and they don't have to fake it.

#### Classwork and Homework Assignments

Student written work is carefully read and, as described in the section on assessment records, above, carefully recorded, and samples kept as individual student records. During the observation period no papers were returned or filed without written comments and all objective assignments were scored with number or percentage right or wrong. Student compositions received evaluative comments one to eight sentences in length. These invariably contained encouraging remarks, noting progress from the previous assignments. Critical comments ranged from noting lack of care with proofreading or penmanship to a too short or superficial job. In math class Clint worked from problems on worksheets or on the blackboard. He collected problem sheets on the average of twice a week, sometimes announcing that this particular work would be collected and sometimes not.

Written class and homework assignments were either dittoed handout worksheets or teacher-composed. Many of the math problems and much of the language arts work on basic grammar, punctuation, and vocabulary were drawn

from the worksheet folders. However, both math problems and vocabulary and grammar problems were also generated by the teachers. . Because Clint and Brenda try to associate the work with ongoing curriculum themes as much as possible, they construct real life problems whenever they can. For example, various things in the classroom were measured as part of the basic geometry instruction. And, a few days before the beach trip, a long list of vocabulary and spelling words was dictated that included various oceanography terms, as well as a list of all the items from home that the students were to bring with them on the trip.

Writing assignments are all teacher-created, some planned as a regular part of each year's curriculum and some generated spontaneously, in response to a class activity or student interest or to meet a particular pedagogical need that the teachers have determined. As an example of a planned activity, the class field trip to the state legislature and governor's office evoked a thank you letter assignment. Each student wrote to thank the governor for visiting with them and to pursue some line of questioning that came out in the meeting, or that they felt should have been discussed.

A more spontaneous assignment was created through brief consultation between the two teachers when a language arts period suddenly became available due to cancellation of rehearsal for the fifth-sixth grade musical play. It was the next-to-last week of school and there had still been some format problems with the last set of letter-writing exercises, so they required a thank you letter to someone who had helped the student through the sixth grade. This assignment, created on the spot, was so successful that it was selected as the writing sample to be sent home with the final evaluations.

At the other extreme of planning, the sixes undertake a series of major multi-media projects during the year, always including a biography project, a physiology project, a study of the local city, and the oceanography project associated with the beach trip. For the oceanography project, the major work underway during the observation period, even the individual student topics remain largely the same from year to year. Assignment and assessment forms are carried over as well. These projects are componentialized, the students first reporting their bibliographies, then an outline of their paper, and eventually first through final drafts of their report text. Once the pre-final draft of the written report is approved, the students adapt it for their oral presentation. At the same time they develop their visual aids, a mural to illustrate their oral presentation and an illustrated cover for their report, as well as illustrated covers for their special trip notebook. Each of these components is carefully assessed and recorded.

Brenda and Clint are completely open about their assessment activities. They take considerable pains to explain to the students when and on what they will be assessed. For example, Brenda carefully explains whether a particular writing assignment is required to be in final or draft form. Students usually ask a number of clarification questions every time an assignment is announced. A questioner who contributes to general understanding of the assignment receives strong, immediate positive reinforcement: The teacher

states that this is a good question and calls the class' attention to it, often requesting that it be repeated if some students have already turned their attention to the work at hand.

Many comments and cues inform the students that assessment is taking or will take place. Film and story sessions offer one example. While Brenda reads to the class or a film runs, Clint takes notes, demonstrating that there will be a question session at the end. Often he speaks over the film, pointing out this or that important point. Students are instructed to take notes during some films and presentations and they are so attuned to the assessment requirements that they frequently stop the speaker or request re-viewing of a film clip, in order to get their notes straight. Many times these requests directly anticipate Clint's comment to note that very section.

The assessment process for the beach trip and oceanography project is complex and multi-faceted. It was explained to the students several times over during my observation period. Brenda detailed just which components were included in a complete project and in what sequence and when they were to be completed and checked. She prepared and distributed an outline of all the parts of and dates for the work that the students were to use to monitor their progress. When interim due dates were missed, she requested that they refer to their check list while meeting with her about their work. Both Clint and Brenda offered comments about quality of the work as it progressed, noting the overall progress of the wall murals on a daily basis and calling the attention of the class to murals that were particularly well done or on which notable progress had been made. In addition to extended comments on all the written components of project, a summary assessment form was completed for each student. (See Appendix.) As announced beforehand, Brenda completed these while listening to each of the students' oral presentations. These general project assessments were then immediately available for the students to study.

The experiential component of the oceanography project was assessed as thoroughly as the academic components. As they well knew, the students were to be assessed on their general behavior during the trip itself, specifically on the extent to which they were able to learn from the activities there and their maturity as responsible group members. Numerous references were made throughout the preparation period to the need that Brenda and Clint would have for quiet and control during the trip back from the beach, for they would be jointly composing their assessments of each student. In addition, the students were informed that their trip counsellors' evaluations would be taken into account. A semi-serious awards ceremony the last night at the beach completed the trip assessments. Each student was given a title that reflected some aspect of their week's experience, for example "best kitchen aide", "martyred sufferer", "dressy camper", "I lost it". Some of these appellations appear in the final evaluations as indications of level of personal and social development.

#### Recitation

Students in the sixth are required to perform publically in a variety of ways. However, with the exception of attention check questions after films or story-readings, they are rarely asked to answer questions unless they are



prepared. It is routine to take volunteers in response to class- or homework assignment checks. If the teachers go around the circle, no negative comment (verbal or non-verbal) is attached to a refusal to answer. Students are not expected to provide any excuse for failure to know an answer to, for example, one of the math homework questions. The teacher simply passes on to the next student. Clint notes that he never wants to "embarrass" individuals this way. There may be a variety of reasons why a student has not completed a specific assignment on time; the recitation period is not the proper venue for exploring that question. Repeated failure to complete work in a timely fashion is, however, taken up with the student in a one-on-one discussion whose seriousness is lost on none of the sixes.

Public performance is required in reading groups and in oral reports and recitations. Ability to speak before the group is an important skill in the eyes of the sixth grade teachers. They assess it, however, when the students have had the chance to prepare for a specific assignment, rather than in on-the-spot questioning. This is consistent with their stated philosophy of having students take responsibility for themselves: asking questions in a testing fashion only undermines their self-esteem and level of trust; asking them to prepare for a recitation, on the other hand, offers them an opportunity to take the responsibility for their performance.

Reading circles are assessed by Brenda, evaluating factors that are also part of her, and Clint's, responses to and corrections of oral reading performance. Her assessments are recorded immediately after the reading session and, more than once during the observation period, she declined to take individual student questions at that time, citing that she was completing assessments of the preceding session. Again, the process and the criteria are well known to the class.

Students are fully aware that oral reports and prepared recitations are important occasions for assessment. In the oceanography project, the teachers reminded the students several times that the oral presentation was every bit as important as the written report.

Another assignment during the observation period was memorization and recitation of a poem of the students' choosing. Clint and Brenda told the class when and how to prepare and made explicit note of the fact that presentation style, not just accurate memorization were important. They likened these skills to those that had been developing in the oral reading groups over the year and assured the students that everyone in the class could, if they chose, succeed well at this task. Brenda explained that she had required a hand-written copy of the poem before their oral performance because she would be making notes on their copy as they read, citing errors and prompts and "delivery". The students watched the assessment process proceed, just as described, and, completing their recitation, were given their copy back, with Brenda's annotations, to be later returned to her "in" box for recording.



## Observation and Interaction

Ongoing observation of student behavior is the fundament of assessment in the sixth. As Clint stated in the pre-study interview, "I'm always assessing them." Brenda and Clint place great importance on a wide variety of student characteristics, many of which cannot be assessed with traditional classroom evaluation measurements such as testing. For them, responsibility is as important as reading and confidence as crucial to math as multiplication. It is not necessarily the high achievers who are the most attractive children to them. Clint put it this way: "...pure academic smarts are just one trick that people have. And of all the people you'd rather be, pick one..."

The teachers constantly observe even subtle student behaviors, interactions among the students as well as their own conversations with them. They are discussed at length between the two; often they review the day after the students have left, filling each other in on student behaviors that occurred when the other was elsewhere. Their memory of the minutiae of student conversations and actions is prodigious.

The observational process is no secret to the students. The teachers frequently comment to the group on how well they have performed, academically and interpersonally, at the end of a class session. Reading circles are an obvious and outstanding example of assessment based on observation. The teachers discuss the oral reading performances of the groups they have conducted on a nearly daily basis. They also switch back and forth with one another, so that each has ample opportunity to observe all the students. And, in addition to the verbal assessments that follow many of the individual turns at reading and the group assessment that the teachers often provide at the end of the session, Brenda keeps occasional, but conspicuous individual records of student reading progress. The task of the reading group is oral reading and comprehension improvement, but assessment is an integrated part of the instruction.

Each day, Clint offers a verbal assessment of how the day has gone to the assembled class. This takes place before afternoon dismissal, if time allows, or first thing in the morning and is based on his and Brenda's observation of class process. On a typical afternoon the wrap up assessment had the following elements:

- o Call for attention.
- o The classroom day has gone fairly well overall. Pretty productive lessons and also homeroom groups were on task.
- o Review of the morning's Lower School assembly. Behavior of group was good, orderly, a model for the little kids. Note Al's good idea about how to set up the assembly.
- o Comment on Clint's visits to the 2 music classes held this afternoon. Found that the sixes were doing ok, paying attention. Also pleased to see some enthusiasm building for the upcoming class performance. Notes that the music teacher is a replacement, deserves their support.

- o Shop is being left messy. More consistent cleaning up is required. It would be best to try to clean up collectively, not just look out for your own mess.
- o One problem in the classroom and in other settings is that requests from teachers are not getting quick enough response. "Sometimes it's important that Brenda and I just tell you things and you do it." Explanation of why such arbitrary demands are necessary, e.g., to save time.
- o Brenda asked to add: Points out the number of names on the board with overdue homework. "I want to see that number reduced by half tomorrow."
- o Announcements: Class newsletter to go home with them; be sure to copy math problems from the board.
- o Final reminder to clean up desk area, put up chair before leaving.

Much of the instruction in the sixth is carried out one-on-one between teacher and student. One of Brenda and Clint's goals is to train the students in their process of instruction, i.e., knowing when and how to ask for help and how to receive it appropriately and apply it. These semi- and private interactions also constitute evaluation opportunities and become data for assessments.

The final evaluations reflect a constant monitoring process. Assessments of academic subjects (these sections, too, filled with observational comments) constitute just about half of most of the evaluations. The other half of the teachers' evaluations is devoted to assessments drawn entirely from their observations of and interactions with the students. This includes a section on behavior during the beach trip, a paragraph on PE performance, often a brief assessment of shop work, and one or two paragraphs of general assessment of personal and social growth. The students are closely observed, as this example illustrates:

Dyan had a great time at the beach nestled in with Kathy in their two-person tent. The girls stuck close together on most activities and tried everything with fairly positive attitudes. Dyan was an independent sort of person, as long as Kathy was close by, but also wanted to check in with us often to receive her allotment of affection. She was not one to offer help but accepted work chores as part of the trip. Frank, Dyan's counselor, referred to her as one who "did what work was necessary with a nice and affectionate manner."

In this classroom observational evaluation is a constant concern of both teachers. And their observations become perhaps the primary basis for their formal assessments of the students.

## Functions of Assessment

### Admission

As a private school, Mount Stanton is able to select among its applicants. Potential students submit their school records and are subjected to achievement testing by the school. Parent and student interviews with the principal precede and follow. The most crucial test, however, by both principal's and teachers' report, is a trial day in the class the child expects to enter. The teachers' observational assessment is the decisive information; their decision is final.

Six or eight applicants visited the sixth during the observation period. Each was assigned a student-buddy who was responsible for talking with the newcomer and preparing an introduction for the rest of the class. At the end of the day, the teachers conferred about whether the child had the potential for functioning well in their classroom. In no case did they find it necessary to study the child's school records; observation of his/her behavior in the classroom and brief, informal discussions with the child were evidence enough. The teachers' criteria included demonstrations of academic ability (e.g., in the reading circle), general deportment, willingness to participate in the classroom activities, ability to interact with at least the assigned student-companion, quality and appropriateness of the child's responses to questions about him/herself and interests and how he/she enjoyed the day.

### Expulsion

The classroom teachers are also the primary decisionmakers in expulsion of students. This has rarely occurred in the Lower School, but last fall Brenda and Clint refused to "run" one boy whose behavior was too disruptive to the class. He failed to get to his various classes on time, talked and acted out of turn in the homeroom, and finally destroyed property and stole from Brenda's purse. The theft enabled Clint and Brenda to have the boy summarily dismissed from the school, but they had already decided to request his removal from their classroom because of his unacceptable interactional behavior and had informed his parents that they would not keep him. He was, in Clint's words, "immature, inappropriate, just not a sixth grader".

### Remediation

It is not uncommon for Mount Stanton students to repeat a grade during their Lower or Middle School years. The sixth had one repeating student this year and at least one student's parents came in to discuss the appropriateness of having their child remain with Brenda and Clint for a second try. The decision last summer to have Noam come back and the decision this spring not to encourage Mark's parents to return their son to sixth were both based on assessments of personal as well as academic skills.

The decision to retain Noam, the teachers report, was easily reached. They were on the verge of contacting his mother last spring to broach the subject, when she came in and asked to have him kept back. His academic achievement levels were below the class norm and he was "just not ready" for the seventh grade. For Brenda it is important that the student is part of the decision: "I like them to feel it will be good for them. And that they feel good enough about us that they would want to spend another year with us." In Noam's case, once teachers and parent had agreed on the wisdom of repeating the sixth, each approached the boy and discussed the merits of such a move. By summer's end Noam agreed and was willing to come back to the sixth. Brenda emphasized that she respected him for being able to make a difficult choice and that this was a sign of his growing maturity.

I was able to observe the conference between Mark's parents and the teachers on the question of their son's repeating the sixth. Clint and Brenda strongly discouraged them from asking Mark to repeat the grade, because they assessed him as "ready" for the Middle School, despite certain deficiencies. The deficiencies were largely among academic, not personal skills, so they felt he would do better going on. Mark is below the class average in math and reading, but making what his final evaluation characterizes as "solid progress" in all the homeroom subjects. He is, according to the evaluation, "able to take care of himself", a "serious student", and "a loyal person and strong member of his class." Mark is the smallest boy in the class and physically young-looking--a child, not an adolescent, as are some of his classmates.

During the parent conference Brenda brought out her assessment record files and offered samples of Mark's work from the beginning and end of the year as illustrations of how far he had come. The teachers clearly communicated their confidence in Mark to his parents; the latter went away reassured and prepared to work with Mark over the summer to better prepare him for the Middle School. In later discussion, Clint and Brenda described their decision about Mark this way:

Clint: He's not a good student, but he's on his way. He thinks that (seventh grade) way.

Brenda: It (retention) would be a disaster.

Clint: He's no longer a little child. I don't care what his academic standing is, it doesn't make any difference how much he's learned by now.

Brenda: Or how big he is.

Clint: He walks bigger than he is.

Mark was not the only student whom Brenda and Clint contemplated as a possible candidate for retention this year. As in Mark's case, with two girls it was personal and social skills, not academic achievement that they weighed most seriously. If it were not for her very high academic standing, they



would have liked to suggest that Suzanna stay another year. She is one of the youngest in the group, small, quiet, and, from their observation, very immature for Middle School. She had recently written an essay about having no friends and not knowing how to make them. In Brenda's opinion, she doesn't try very hard either. A second girl, Ginger, is also physically young and her academic skills are more precarious. Both teachers predict that she may well repeat a grade before going on to the Upper School. "Right now", says Brenda, "she needs to stay with her group. She's not ready to appreciate a chance to catch up and get ahead. But later she'll see what an advantage it can be."

While the teachers only occasionally recommend that a student repeat a grade, they often suggest remedial work. At Mount Stanton most of the parents are willing and able to contribute substantial effort to help their children with homework and extra assignments. In one case this past year, a father met weekly with Clint and Brenda during the fall, in order to be better able to assist his son. As with Mark's parents, Brenda provided materials for the parents to use to work through special problems with their children.

For many students the final evaluations contained recommendations for summer work at home, and, in two cases, enrollment in a structured summer remedial studies program. Reading practice was most frequently suggested: Twelve students' parents were asked to spend time listening to stories over the summer. For seven students summer math work was recommended, with the specific skill noted. Two students, for example, did not have full mastery of the multiplication tables and memorization work was suggested. Tennis or some other active sport was recommended for three students whose physical coordination is poor. And parents of one student whom Clint and Brenda had difficulty engaging in conversation were asked to emphasize verbal interaction with their child.

### Sizing Up

Although Brenda and Clint are provided with the results of the fifth grade achievement test, as well as school records and test scores for students new to Mount Stanton, the sixth grade teachers prefer to size up their new students through classroom performance before they look at any documentation. It is several weeks into the new term before they examine the test scores and then it is to check against their own sizing up. In general, there is a high level of congruence, they report. Discrepancies that do occur are usually higher performance on the self-designed test than the standardized test results would predict. This confirms the teachers' distrust of high-stress testing situation results.

In fact, Clint and Brenda both report that they go out of their way not to get to know the students in the fourth and fifth grades but in order see them all fresh and without preconceptions when they arrive in the sixth grade classroom. Clint cites Ulrich as an example: "He's widely regarded as someone who skates around things, is a cut-up. I'd just as soon not know that that's how his fourth grade teachers look at him."

They plan a series of activities in the first weeks of the fall term to size up the students' academic and personal skills. Clint has developed his own math test which is given the first week of school. It covers material that should be mastered by the sixth and also material he plans to cover. This initial math assessment also includes careful observation of the students' anxiety and self-confidence levels. He and Brenda check the test papers for clarity, neatness, and carefulness, as well as accuracy.

The reading circles start immediately, and Brenda and Clint begin their assessment of the students' oral reading and comprehension skills. They select the first books according to their observations. The students have already been grouped for French, according to background in the language, so the teachers are also monitoring whether some of the new Mount Stanton students have advanced language arts skills, despite their French placement.

The first major language arts project serves to size up a wide variety of skills. Each student is assigned a classmate whom he she is to interview and then introduce to the class. The oral presentation must include a biography of the classmate and some interesting things about her/him and be accompanied by a portrait of the subject that the student has drawn. A written version of the introduction is also submitted. This assignment works well to create a sense of group cohesion, the teachers find, since there are always a number of newcomers to the class. It allows them to size up the students' writing and speaking skills and also their ability to get to know another person, their interest in others, and how insightful they are about other people. Brenda explains: "You find out who isn't thinking, who has trouble thinking. Somebody will ask five questions, get 'yes' or 'no' and think they're done." Brenda specifically points out that the drawing of the classmate is often a revealing factor. They range from stick figures to carefully executed and often striking likenesses: "A caricature, a stick figure, or a joke is, 'I can't do it,' sort of." The process of presenting another person in art tells a lot about the children's self-esteem and confidence, as well as their depth of understanding of others. Further, the teachers note that they can learn a great deal about the students' attitudes toward learning, as they watch later presenters revising and redoing their work to follow good models they have already seen.

Brenda and Clint also assist the seventh grade teachers in sizing up their outgoing class. Because the Middle School format has students meeting with a different teacher for every class, the seventh grade teachers, as Brenda puts it, "don't have the time to get to really know them all" early in the term. Thus, they must rely upon achievement test scores and student records far more than the Lower School teachers. But Brenda and Clint are given the opportunity to go over the placements before they are finalized and often make suggestions. They have sent several notes to teachers in the Middle School about certain students, for example noting a learning problem that is not apparent in a bright child and suggesting that some students should be grouped together and others not.

## Grouping

Achievement grouping is a controversial practice in the Lower School. Brenda and Clint personally regard it as counter-productive. Yet the students are grouped for two levels of French--Clint calls it "experience grouping"--and, as a result, are also grouped for the reading classes that take place during the French session on alternate days. Most of the homeroom instruction is structured so that it involves only a portion of the 42-student class. Reading breaks the class into four small sections; math into two groups of about twenty; and the two afternoon sessions in which the students alternate among science, art, music, shop, and homeroom engage a third of the students at a time. With the exception of the two French levels, all these groupings are carefully constructed to combine, not separate, students of varying background and ability.

For French (and therefore reading) the class is divided into two equal groups: those who have done well in the language and those who have done poorly, along with students new to the school's language program. This effectively puts most of the students with strong language arts ability into one of the halves, although some of the incoming students have strong skills in language arts, despite their French I placement. This Brenda and Clint regard as a critical advantage, since there remain some good reading models in the lower language arts groups. The further division of the two halves of the class into small reading circles is entirely at the students' discretion. Two different books are offered and the students select which they would like to read. When a pair of books is completed, the two circles assemble, discuss new selections, and reconstitute as new groups.

Math sections, too, are consciously structured to avoid ability or achievement grouping. During the observation period there was a small, special math group that had recently been formed of students who had been absent several weeks on a school trip. They were temporarily getting an extra session, so that they could catch up. Otherwise, the two math sections were divided in an apparently random fashion. Clint says he gives his sizing-up math test "just to see where the kids are", but does not use the scores to structure the groupings. Grouping kids is "too bossy". For each assignment, he suggests that the students look at their scores and consider whether they need more work on that skill or not. With Brenda's assistance, he subdivides the class into reviewers and those who are going on to something else, telling them, "If you think that you want to work on this stay here; if you don't, go do something else." Typically, Brenda works with those going on to the new skill and Clint works individually with those who are reviewing.

The teachers work hard to defuse the stigma of seeking more help. Clint often cites his own school years: "If you can't do it--and I couldn't at 12--come on over and we'll bang it out." Or, as he retells it to them, "I couldn't do this in the sixth grade and I still grew up to be a math."



teacher." By late in the year, Brenda says, they are able to realize the self-responsible behavior they seek:

It works, it really does. It's amazing. If we take the papers home and correct them and enter them in the book and just sort of mentally divide them up. So the next day you go back in and you actually don't even have to pass the papers back. The kids know if they've done well or not. They know if they know it and know if they don't. And if you give them the choice of "If you know it stay here and if you don't go over there" they will almost always put themselves the same way we would.

Indeed, this behavior could be easily observed during the math classes. When the class divided itself for problem work, those who had done well shifted to the "new work" section of the room and a cluster of help-seekers would form around the review instructor. Some students would go to Brenda and work for a while, then just say, "I'm going to see Clint now" and go back to get the review. Clint and Brenda believe that a non-graded environment is crucial for developing this level of openness in the students.

Even when groupings do not work out well, Brenda and Clint are reluctant to alter them. One example came up during the observation period. Noam, their repeating sixth grader, had become a discipline problem for the art teacher, cutting up with one of his buddies. The art teacher requested that Noam be moved to another of the afternoon groupings, so that he would be separated from the other boy. Clint and Brenda considered this question in their after-school conversation. The primary reason they decided not to fulfill the request was that Noam "really needs to take responsibility for himself. This (forced change) would just make it our problem, not his." To impose such external discipline would be directly contrary to their pedagogical philosophy. The teachers chose instead to have a serious talk with Noam, pointing out to him the effects of his behavior on himself and others.

### Diagnosis

Diagnosis of individual student academic progress is derived from the written work turned in as homework and class assignments, from in-class performance, and from one-on-one work with the students. Assignments are tailored to the individual student in certain cases. Some will be asked to redo a paper for a third or fourth time, if it still shows weaknesses. Others are asked to do an extra sheet of math problems in an area in which they are still having trouble. In most cases the teachers suggest, rather than command. An oceanography mural, for example, was considered done by the student, but Brenda asked if it were sufficient to help the student explain all the points in his talk. The student deliberated and chose to expand his work. During in-class work corrections and diagnosis occur almost exclusively in one-on-one sessions between student and teacher. Brenda also calls students to her desk to discuss written work, using these explorations to form diagnosis of that student, sometimes gaining impressions of misunderstandings that may be more general to the class as a whole.



The teachers use homework and in-class checks to diagnose the progress of the class. The math homework, for example, is collected about twice a week, for, according to Clint, "You wouldn't know what they're up to if you didn't." They often stop while giving an assignment to see if there are questions and use the questions to clarify the assignment, as well as make further instructional comments. They also stop to ask the group, in general, how they are doing. In math, for instance, Clint pauses after a board problem or an explanation and requests, "Raise you hand if you know how to do this." At such a point he may decide to review or go on, depending on the percentage, or divide the class into new and review sections.

Diagnosis is an ongoing process, the teachers say, an integral part of their instructional work. It is also a judgement call, something that you can only learn by experience and which requires that you really know the students well. Just after a math class I asked Clint to reflect on one of his diagnostic activities. The group had been working on problems relating to area and volume. A number of students used "square" for the measure, when "cubic" was called for. In some cases Clint stopped and corrected the student, in others not. Why the distinction, I asked. "I know the kids", was his response. "Sometimes it's just sloppiness or carelessness, but they know the concept. In some of the others it tells me they just don't have it yet." These kinds of performance errors, he went on to say, are important information for making instructional decisions.

#### Feedback to Parents

Parent feedback is a structured part of Lower School work, including two written evaluations (serving in the place of grade reports) each year, individual parent conferences, parent information nights, and a monthly class newsletter. The evaluations and conferences have been described above. Brenda's files of student written work are available for and used during parent conferences. These files are sent home with the students at the end of the year. No parent nights took place during the study period.

Two examples of the sixth grade newsletter are included in the Appendix. The first fall newsletter outlines Brenda and Clint's instructional and assessment practices. Subsequent issues include—along with various announcements and updates—information on the progress of the class in the homeroom subjects. Parents are informed about the specific skills their children are expected to have mastered.

In the fall, and occasionally throughout the year, parent review of homework is required. In order to impress both the students and the parents with the importance of regular and prompt homework completion, notes are sent home that require parents' signatures, affirming that they have looked at the completed assignments. Brenda also occasionally writes short notes on "signature sheets" outlining how the student is doing. This strategy is used when individuals or the group are irresponsible about their work. Brenda reports that many of these come back with more than a signature. Some parents write notes, many thanking the teachers for their careful monitoring of student effort and some detailing what work they do with their children at

home. These comments are very helpful in deciding how to proceed with an individual student, Clint and Brenda report. It's valuable to know "if we're doing the job alone."

#### Feedback to Students: Control and Motivation

While plenty of control statements occur during the course of a day in the sixth, relatively few of them are directly related to assessments. As with any group of 40 12-year-olds, intermittent calls for quiet, attention, and order are required. Those that are assessment-related are overwhelming lectures on timeliness—failure to be punctual for classes or to get down to work immediately upon arrival and failure to complete tasks on time. Brenda and Clint use public listing of those with overdue assignments and, eventually, recess period and after school detention to discipline habitual offenders.

The only other occasions of "balling out"—as Clint characterizes his more aggressive lectures—occurred when students violated class norms for politeness to teachers and other students. These included talking when a student or teacher was presenting, pursuing some private game or interaction with a neighbor during presentations, and, once, two students walked out to get a drink while Brenda was reading a story aloud to the class.

In cases in which Brenda and Clint deemed a serious infraction had occurred, they signalled the importance of the impending interaction by several cues that were apparently readily interpreted by the class. Clint rarely raised his voice, but when he did, attention and quiet were instant. All his serious "ballings out" took place from the same location at which he stands for class announcements and group instructions—the teachers' speaking chair. (See Figure One.) If he started a disciplinary talk and even began moving to that location, the effect was the same as raising his voice—quick attention. Brenda did not ordinarily use the teachers' speaking chair as her forum, generally remaining in the area of her desk. On one occasion, when she was very put out that few students had heeded her demand that a specific assignment be completed, she moved to the speakers' chair to announce, in her usual quiet voice, how irresponsible the class was being. The students recognized the use of that space by Brenda as unusual and important. They comprehended the seriousness of her reprimand and responded unusually quickly, getting down to work on the assignment.

Generally, assessment feedback is used to motivate students, especially to motivate them to put out effort, to be self-monitoring, to feel successful, and to take pride in their good work. For example, math homework is always scored, so that students "know where they stand", Clint explains. If they get the percentage or number right, they can easily see if they need more help and it is then their responsibility to seek it. Whenever a student performs in front of the group, verbal assessment comments are made by the teachers. Often they explicitly state that "That's how it should be done", i.e., that this is a model for others to follow. Some of the teachers' proudest moments came when students voluntarily undertook to redo their assignments after seeing a classmate do a better job than they had planned to do.

One of the goals that Brenda and Clint cite for the oral reading circles is that "everyone feels successful". There are so many components to reading aloud that most of the students can do a good job of one or some of them, e.g., knowing lots of the words, speaking in a humorous or animated fashion, being able to speak in dialect, remembering the background of the story. The teachers try to comment on successes by every student in the course of their discussions with the class. For one student whose academic skills were below average, the teachers pointed out the excellence of his mural drawing many times. For several others shop projects were commented on positively or displayed to the class. Yet others appeared in the afternoon wrap-up as models for success from their PE performance. Specific instances in which a student had been observed helping or complimenting a classmate, making a clever statement to another student, or asking a good question were all recalled at the end of the day for the class to appreciate and learn from.

Students in the sixth are encouraged to plan their own work. Initiative in this area is immediately rewarded and often recalled in class wrap-up and in final evaluations. Several interesting examples were observed. Larry and Patti had failed to complete their oceanography wall mural the afternoon before their scheduled presentation. Brenda's instructions were that the mural must be done first, since it was to illustrate the talk. The two students came to her in their homeroom period and negotiated to come early to school the next day to work on it and to have their talk rescheduled from the morning to the afternoon. They would stay in through recess period and come back early from lunch to finish up. Brenda agreed and concluded the conference with a discussion of how they had got themselves in this difficult position and what they had learned from the process.

When, the last week of classes, the students were assigned a thank-you letter and required—in very vociferous terms—to get it in by the end of the day, Brenda accepted only one extension. With the exception of Mark, everyone who did not complete the letter had to stay after school. Mark is a below-average writer, but a serious and hard-working student. He came to Brenda during the latter part of the day in great distress. He had been working on his letter, but was clearly not going to finish. He asked if he could take it home and complete it. Brenda describes her decision:

Mark is a slow and difficult writer. He simply would not have been able to cope if he'd had to turn it in today. He knew that, at this point in the year, we were going to have to have a job that takes longer, instead of a quick, shoddy one. We've really been working on writing with him and he's come so far. His last book report was terrific—sensitive and well-written. I wouldn't have traded that for all the on-times that are possible.

Ideally, the sixes function in peer-like relationships with their teachers. Willingness and ability to negotiate about their work is, for Brenda and Clint, one of the strongest indicators that they have succeeded with their students. Clint interpreted for me a conversation he had with Tim,

who had returned from an extended illness to be elected class president. The presidency required that he do quite a bit of organizational work for the sixth's final assembly. Tim came to Clint and asked to be excused from a project. It was agreed.

He came around saying, "I've got a lot to do as class president and I can't do the city project." I said, "Fine. What you're doing is just as important." Now, is that sensible? Two reasonable people just talking to each other....That's the difference. We're not a policeman and we're not a heavy.

These are among the most successful and rewarding moments for Clint and Brenda. They see their students taking the work seriously and wanting to complete it well, rather than simply meeting the minimal requirements of the assignment. They value this positive and self-responsible attitude to learning higher than completion of any specific task or attainment of any specific skill or knowledge. Rather than "policing" student work, at the best of times they are able to function as negotiators and facilitators as students undertake, under their guidance, endeavors which become the students' own.



## Students as Assessors

Self-assessment is a stated goal for the sixth grade teachers. Brenda and Clint assess their students, and themselves, in terms success on this question. It is communicated to the students in a variety of ways, formally and informally: Level of self-responsibility and self-monitoring is the most frequently noted behavioral characteristic in the students' final evaluations; students are constantly reaffirmed for taking responsibility and initiative in class; instances of student self-responsibility are pointed out to the rest of the class as models for behavior.

One of the last major writing assignments for the sixes offers the teachers an index of their success in making the students self-assessors. Each student is required to complete a self-evaluation. (See Appendix for self-evaluation assignment sheet.) Students are asked to evaluate their subject matter classes, their trips, and their projects; to describe what has been their greatest success and failure in the sixth; and to consider goals for the coming year. Table Eight displays the comments made in eighteen self-evaluations. The self-evaluations ranged from one to seven pages in length and addressed an average of over twelve topics. They vary greatly in detail and in tone.

Every self-evaluation contained at least some comment on progress in a subject matter area. Of the 139 comments on academic skill level, all but 12 comments reported improvement. The remainder were expressions of disappointment that either progress had not been as great as the student had hoped or self-chastisement that the student had been unable to apply her/himself sufficiently to the subject. The homeroom subjects, writing, reading, and math, were of primary concern to the students, evoking almost three-fourths of the comments (82). The specific skills mentioned for those three subjects closely parallel those noted by the teachers in the final evaluations. (See Tables Two, Four, and Six, above.)

Remarks on academically-related skills and on goals for the seventh grade also dovetail with the teachers' assessment topics. The first paragraph in the teachers' final evaluation assesses the students' behavior on the beach trip; the most common comment by the students is the extent to which they were able to learn from trip experiences (13 comments, all positive). "Pride in work" is equal to "task completion" in the number of comments (12 each) and the students also share their teachers' values about quality of work and risk-taking. (See Tables One and Seven, above.) It is not only standards for punctuality, task completion, and effort that Brenda and Clint have communicated to their students. The sixes have come to understand that more complex and abstract skills are expected of them and that those personal characteristics are important for their academic progress.

For 17 students the self-evaluations contain enough material to enable one-on-one comparison with that student's final evaluation. In only two cases is there substantial disagreement between the evaluation the teachers were then in the process of preparing and the student's self-assessment. In one of

TABLE EIGHT

Student Self-Evaluations

| <u>Subject Matter Skills</u> |     | <u>Academically<br/>Related Skills</u> | <u>Goals for Seventh Grade</u> |
|------------------------------|-----|----------------------------------------|--------------------------------|
| Writing                      | 34  | Learn from trips                       | 13                             |
| Spelling                     | 9   | Task completion                        | 12                             |
| Penmanship                   | 5   | Pride in work                          | 12                             |
| Grammar                      | 3   | Quality of work                        | 9                              |
| Reading                      | 27  | Punctuality                            | 6                              |
| Interest                     | 3   | Try new things                         | 5                              |
| Phrasing                     | 2   | Attitude to                            | 3                              |
| Comprehension                | 2   | learning                               | 3                              |
| Scanning/<br>wording         | 2   | Make friends                           | 3                              |
| Math                         | 21  | Effort                                 | 2                              |
| Decimals                     | 3   | Talk in class                          | 2                              |
| Fractions                    | 2   | Classroom behavior                     | 1                              |
| Percentage                   | 2   |                                        | 68                             |
| Multiplication               | 2   |                                        |                                |
| Reports                      | 17  |                                        |                                |
| Vocabulary                   | 2   |                                        |                                |
| Other subjects <sup>1</sup>  | 38  |                                        |                                |
|                              | 139 |                                        |                                |
|                              |     |                                        | 27                             |

n = 18

<sup>1</sup>French (12), shop (8), PE (7), art (4), music (3), Spanish (1), geography (1).  
<sup>2</sup>Shop, PE, science.

these exceptional cases the student reports a far lower level of self-esteem than do the teachers. In the other, the student is not as satisfied with her progress in various subject areas as are her teachers. In all 17 cases, there is high congruence in choice of subject for comment. Almost all the specific topics the students select for comment are among those cited in the teachers' more detailed evaluations. Two examples will illustrate how fully many sixth grade students are able to assess themselves.

Larry is a good student academically and socially. He writes in his self-evaluation that he has "learned alot (sic) in math, language, and French" and that he has "looked at learning from a different standpoint" this year. Specifically, he reports that he met his goals of improving spelling and writing. He feels he still has some carelessness in math, but has improved his skills. Next year he wants to be more accurate, too. Larry enjoyed his oceanography project. It was "one of the funnest (sic) reports" he has ever done and he feels good about learning how to do research as well as his experience working with his partner on the mural part of the project. He was able to use his research on crabs, he says, to learn things at the beach and had fun and learned a great deal there. He does worry that he has the habit of procrastinating on large projects and sets more steady work as a goal for next year. Overall, Larry reports that he is "most satisfied with all that I have learned this year."

Brenda and Clint's final report to Larry and his parents highlights many of the same accomplishments. They observe that the beach trip "was a high point of Larry's sixth grade year." He is a serious student, learned a lot on the trip, but also took time to have fun and enjoy his classmates and the counselors. He should be satisfied with his progress in all areas, social and academic, they report, for he is a leader and "most definitely respected and followed by all who know him." Larry's self-criticisms would come as no surprise to his teachers. He is "a responsible student" and good academic work "is a high priority in his life....He sets high standards for himself and meets them." Specifically, they note that he "paced himself pretty well on major reports", an area in which Larry himself would like more improvement. And they agree that his research skills have increased: "The small research classes were useful to him." They also note that he "gives credit to those who provide help," as, in fact, Larry does give credit to his co-worker on the crab mural. (Larry likewise takes space in his self-evaluation to thank his homeroom teachers, writing "I've learned so much because of the teachers.") In his homeroom subjects Larry and his teachers agree that he has improved in spelling and the teachers report, as does Larry, that his writing is good. Larry feels he is not yet careful enough in math, and the teachers report that, while he is "careful, methodical", he "does get frustrated easily." They concur with Larry's insight that he has begun to see his education "from a different standpoint": this student, in his teachers' view has developed "a mature understanding" of his educational life.

Kathy's self-evaluation and Brenda and Clint's final evaluation of her show equally striking parallels. For example, Kathy writes that "My reading has stayed the same", while her teachers report, "Kathy's reading skills have remained pretty constant over the year." Both student and teachers agree that

Kathy's reading skills were and remain above average in the class. Both express concern over her physical condition, Kathy writing that it was, "one goal that didn't work out" and her teachers commenting that, while she is making progress, they would recommend some physical activity for her over the summer. Writing is, from Kathy's perspective, "always a fun subject" and her skills are "at a good healthy level". From the teachers' point of view, "Writing is easy for Kathy and perhaps her strongest suit." And, in the area of social and personal growth, teachers and student alike suggest that making new friends should be a goal for Kathy in the upcoming school year.

This year's sixes were also set a second task that gives some insight into their understanding of their teachers' assessment process and educational goals. Each was asked to create a list of suggestions for the incoming sixth graders on "how to succeed in the sixth". Table Nine details their responses. The 18 lists available included one to 12 items, with aspects of relations to the teachers most frequently mentioned (18 of 168 comments). Some of the lists were all humorous; most included a few jokes at the teachers' expense (e.g., "Don't get Clint mad"). Here, too, the teachers' values of "be nice", "be responsible for yourself", "try new things", "keep trying" show up as well as the more obvious priorities of task completion and punctuality, although the latter are by far most frequently reinforced in the classroom.

When one of the students asked me to explain more to him about what I was doing in his classroom, he responded to my explanation of classroom performance assessment research by sharing his own ideas about assessment. Tim had spent his first years in public school, he told me, and he preferred Mount Stanton because there are no grades.

Our evaluations are better than grades, because they tell you what's wrong. I talk too much. In public school I would just get a "D". Math, it's not a problem, but with math I talk a lot, so I'd get a "D". But I know how to do it real well.

The same, he says, is true of reading, where he would get a low grade because he talks, but "It (a low grade) wouldn't mean anything." A grade "doesn't tell me what to do" to make it better or to improve in substantial ways. Usually he agrees with what Brenda and Clint say on their evaluations, even the negative things. On the times when he didn't agree, "I told my parents I didn't do that and then I just went somewhere and read or something."

Success, says Clint, is "When you get the kids not to think they have to do things, but when they feel freer and more independent." Brenda agrees. Her day is made when students take initiative to think about an assignment, when they ask themselves:

"...What is the assignment? What should I do with it?" The whole attitude—what kind of creativity they throw into it....You know, there are a lot of things that they come up with that would be OK (as



TABLE NINE

Student Assignment "How to Succeed in the Sixth"

|                         |            |
|-------------------------|------------|
| Relationship            |            |
| to teacher              | 18         |
| Don't fear              |            |
| them                    | 6          |
| Don't get Clint         |            |
| mad (facetious)         | 5          |
| Respect &               |            |
| use them                | 4          |
| Be affectionate         | 3          |
| Complete homework       | 14         |
| Be nice/kind/polite     | 13         |
| Be punctual             | 10         |
| Don't be sloppy         | 9          |
| Give full effort        | 9          |
| Don't cut up            | 8          |
| Be flexible/patient/    |            |
| good sport              | 8          |
| Have fun                | 8          |
| Be prepared to          |            |
| work hard               | 8          |
| Be self-responsible     | 7          |
| Try new things/         |            |
| take risks              | 6          |
| Keep positive           |            |
| attitude/keep trying    | 6          |
| Accept criticism        | 5          |
| Develop confidence/     |            |
| self-esteem             | 5          |
| Don't swear             | 4          |
| Don't talk out of       |            |
| turn                    | 3          |
| Set example for         |            |
| younger grades          | 3          |
| Listen in class         | 3          |
| Be organized/consistent | 3          |
| Be direct/honest        | 2          |
| Make friends            | 2          |
| Follow directions       | 1          |
| Comments on other       |            |
| subjects <sup>1</sup>   | 13         |
|                         | <u>168</u> |

n = 22

<sup>1</sup>French (7), shop (5), art (1).

a form of the assignment). Like today, when Emily asked, "What about making a card instead of a (thank you) letter?" I told her, "If it's an interesting card." [Elaine:] "I could decorate it with flowers all around." That was real nice; she was thinking.

Failure, Clint adds, is when the students don't take that kind of responsibility and initiative: "I'm always discouraged when somebody says, 'Is this what you want?'" Judging from their students' performance and perceptions, Brenda and Clint have little to be discouraged about.

**APPENDIX**

STUDENT \_\_\_\_\_

TEACHER \_\_\_\_\_

GRADE LEVEL SixthDATE June 1985

\_\_\_\_\_ will probably never forget her \_\_\_\_\_ experience, nor the feeling of success at bagging the biggest clam. She loved everything about the trip: being out of doors, bunking with friends in a tent, and trying all sorts of new activities. Anna is not a bubbly personality but affectionate and appreciative of the friendship she shared with teachers and adults. She supports her peers in a subtle way and is quick to help someone in need. Above all, \_\_\_\_\_ was reliable. If a task was appointed, it was certain to be completed correctly if Anna was involved. \_\_\_\_\_ counselor, found her to be straightforward and a good worker. He noticed that, although she was not outrightly friendly, she was a caring, pleasant individual who worked hard without complaint. We were happy to have Anna along and know she will look back on the experience with fond thoughts.

Sixth grade has been a good year for \_\_\_\_\_ to solidify old study habits and develop new ones. She is a more independent and responsible student than she was last fall. The small research classes were helpful in organizational skills and breaking down large topics. \_\_\_\_\_ paid close attention and used her education wisely.

Writing is easier for \_\_\_\_\_ now and more reflective of her depth of education. Her thoughts show cohesion, are well-written and include her broadening vocabulary. \_\_\_\_\_ likes to write which is the biggest contributing factor in her growth. She pushes herself as others always hoped she would, and she respects the value of writing as a form of self expression. There is a great deal of pride reflected in \_\_\_\_\_ work, which has developed over the year. She is getting better at organizing her own papers, is a solid spell checker and gives penmanship due attention. It has been fun for us to be a part of this significant change in Anna, and we certainly do appreciate her sincere efforts.

The small reading groups have been fun for \_\_\_\_\_. She is beginning to appreciate the value of literature and learns through the experiences of others. \_\_\_\_\_'s favorite selection this year was The Pearl by Steinbeck. Because of the nature of the subject matter and heavy emotion, it is unusual for a sixth grader to actually appreciate Steinbeck's writing. \_\_\_\_\_ also enjoyed the humor in Farley Mowat's The Dog Who Wouldn't Be, another difficult book with a stiff vocabulary. In addition, she completed In Search of the Tenth [1] Crane and Across Five Aprils. \_\_\_\_\_ continues to chide herself for omitting small words or being jerky at times, but she has made good progress. \_\_\_\_\_ her reading this summer, aloud whenever possible.

\_\_\_\_\_ is a better math student than she lets on. She worked hard to become a mathematician rather than just completing problems. Her conceptual knowledge has improved notably, enabling her to use skills with more consistency. \_\_\_\_\_ is methodical, accurate, and in tune with new concepts. She wants to learn all there is and will correct her own mistakes voluntarily as a means to that end. She proofreads her work catching many of her own errors in the process. \_\_\_\_\_ scored rather high on the standardized exams and will be placed in a challenging math group next fall.



Softball has been the game of this sunny spring, and the sixth graders have generally learned to appreciate this venerable old game. [redacted] is beautiful to watch. She enjoys herself, has good skills, and is a valuable team member. She is beginning to understand the intricacies of baseball. She likes competition and gets a lot of mileage—and exercise—out of P.E. periods.

Congratulations to Anna and her go-cart group! Their wooden limousine is now on the road and is something to be proud of. Somebody's car is going to have to move over in a garage somewhere. Nice job!

[redacted] has gained confidence and independence of thought this year. She is certainly her own person and not to be taken lightly. She asks good questions because she wants to know answers. She either agrees or disagrees or is thinking about it, but she'll let you know. There is an honest, shoot-from-the-hip approach to her that we find most refreshing and has made her into the respected and admired classmate she is. She isn't so serious that it gets in the way, and she can laugh at herself which is something she must keep. She has good friends, she works hard, she sets out to do the job, and we are delighted with positive progress socially and academically.

It's been a wonderful year for [redacted]. We hope she drops by to give us a hug when she becomes a well-knowledged Middle Schooler. We'll miss her.

STUDENT \_\_\_\_\_

TEACHER \_\_\_\_\_

GRADE LEVEL SixthDATE June 1985

\_\_\_\_\_ loved and welcomed every aspect and offering of the summer experience. He was generous with his sincerity, although not very well with thank you or affectionate appreciation. He recognized the amount of effort and effort with a trip requested and showed his appreciation through interest and in other ways. \_\_\_\_\_ loved learning how to do new things. From swimming and crabbing to chopping wood and using a pocket knife. During the week \_\_\_\_\_ was closely knit with his friends and liked that life. He was responsible with chores and showed up for everything ready to do his part. \_\_\_\_\_ is not a team-minded, as we would like him to be. He forms impressions too quickly and seldom cares to venture beyond them. It is to his advantage to open to new, broaden his horizons and be a little more aware of what is being done for him. Overall, though, it was a good week, and she will never forget it.

Reading has become a pleasure to \_\_\_\_\_ since he has discovered that reading aloud is not a chore but it was a reward. He has concentrated and worked hard over the week, learned accurately, smoothly and with more expression. He is just transcending from reading words individually to phrasing and scanning ahead, which is part of the process. He still doesn't catch on to the meaning and needs more focus on comprehension, but he is getting better. \_\_\_\_\_'s favorite selection of the year was Old Yeller but he also enjoyed in the Mount Homer of The One Who Would Not Be. In addition, \_\_\_\_\_ isolated In Search of the Goodwill Crane, The Pearl, some short stories in the larger textbook and Qals in the Family, another one of his favorites.

Perhaps because \_\_\_\_\_ is just beginning to enjoy literature, his writing also reflects a borderline discovery approach of the same nature. It doesn't like to write and will be the first to admit the problems which, until lately, have inhibited his written expression. It has taken \_\_\_\_\_ a long time to build up enough confidence to write at all. The attached evaluation is a good sample of \_\_\_\_\_'s writing level. He displays deeper thought, more cohesion and better organization than we have before. His understanding of the process is improving, and he is able to focus on a writing assignment, which is fairly recent. \_\_\_\_\_ enjoyed the \_\_\_\_\_ report and did a good job all around. He pulled from the small research studies and used his time productively. Again, focus and maturity came into play.

\_\_\_\_\_ 's math abilities place him in about the middle of the class. He has a new attitude toward math having learned some relevance through the making of the go-carts. \_\_\_\_\_ moves easily among fractions, decimals and percents for the first time. His basic facts, multiplication, etc., are pretty solid. Story problems are \_\_\_\_\_ 's biggest hurdle at the moment and directly related to his reading ability. He is strongest in understanding concepts and does a pretty good job with computations. \_\_\_\_\_ is prone to careless errors, like not reducing fractions or forgetting to transfer decimals, but he has become more conscientious over recent months. There is a competitive streak in \_\_\_\_\_ which is helping him along, especially because he competes with himself. He wants to do better than he did last time and is feeling noticeably proud of his success.

--continued--



Page Two  
June 1985

GRADE 11/12

P.E. these days is mostly softball. [redacted] can take it or leave it, but he is learning the game and does pretty well. He often forgets his outfit—forgetful of the real nuisance with [redacted]—and there have been more incidents than necessary over it. He doesn't like to change, to remove his jacket—that sort of thing—and he resists doing it. He ought to grow up a bit and go along with fussy teachers. Sometimes we have reasons...

[redacted]'s go-cart may not make it this season but perhaps can be finished over the summer. I'd like to help—some evening I'll show up at your place with a bag full of tools and we'll put it together. He should organize, though, and give me a call this summer. Don't wait!

[redacted] has spent a useful time in his life with us. He is a stronger student, he has made some good friends, and he has grown up a lot. He has strong likes and dislikes, but his judgment is maturing, and he listens to reason more than he did earlier. He is not afraid to pursue a point and has just enough humor to laugh when he finds things aren't quite what he supposed. He possesses a manly nature and strength of character not often seen in people his age.

BEST COPY AVAILABLE

STUDENT \_\_\_\_\_

INSTRUCTOR

Woodshop

COURSE

Woodshop

DATE

1-25-85

In Shop we strive to develop an appreciation for wood and the ability to work with it. The following categories represent the activities and attitudes we consider to be important in this endeavor.

1. Understanding of the tools:  
a) Excellent \_\_\_\_\_ b) Good ☒ c) Average \_\_\_\_\_ d) Poor \_\_\_\_\_
2. Understanding of the materials:  
a) Good \_\_\_\_\_ b) Fair ☒ c) Limited \_\_\_\_\_
3. Expresses ideas through drawing  
a) Eagerly \_\_\_\_\_ b) Willingly \_\_\_\_\_ c) Hesitantly ☒
4. Seeks help and follows directions.  
a) Always \_\_\_\_\_ b) Usually ☒ c) Seldom \_\_\_\_\_
5. Use of time:  
a) Very effective \_\_\_\_\_ b) Usually effective ☒ c) Fair \_\_\_\_\_ d) Poor \_\_\_\_\_
6. Work in a group:  
a) Always works well \_\_\_\_\_ b) Usually works well ☒ c) Works poorly \_\_\_\_\_
7. Initiative, curiosity, perseverance:  
a) Outstanding \_\_\_\_\_ b) Satisfactory ☒ c) Reluctant \_\_\_\_\_
8. Shows pride in workmanship  
a) Always \_\_\_\_\_ b) Usually ☒ c) Sometimes \_\_\_\_\_ d) Seldom \_\_\_\_\_
9. Shows respect for the shop.  
a) Conscientious \_\_\_\_\_ b) Moderate ☒ c) Indifferent \_\_\_\_\_
10. Additional comments: \_\_\_\_\_ spends a lot of time in shop - including recesses, and enjoys it considerably. He has imagination and a good sense of tool. He does need to be more persistent in following his projects through to completion. He's gotten great starts but not enough follow through. Focus more and avoid distractions!

BEST COPY AVAILABLE

Instructor \_\_\_\_\_



Sixth Grade  
Self-Evaluation Assignment  
20 May 1985

EVALUATION

Food for thought:

What are you proud of and feel most satisfied with about this year?

What personal and academic goals have you met?

What efforts fell short?

What should be or are your goals (or plans) for the '85-'86 school year?

Evaluate:

reading, writing, language, French, music, shop, P.E., art

Projects:

geography, physiology, <sup>beach</sup> city study

trips:

Ar, 's cabin, Mexico

Tips, notes, comments, suggestions to and

November 5, 1984

Dear Parents,

We have moved from Halloween, U.N.I.C.E.F., conferences, O.E.S. soccer games, and fractions into the month of Thanksgiving, food drives, rummage sale, Charles Wright and decimals. The sixth grade is running smoothly despite occasional changes in the schedule.

First, we want to say thank you for taking time out of busy schedules for conferences. It is a special time for us when we can meet and discuss individual students.

The trip to the Multnomah County Library was successful in every sense of the word. Students discovered that their research skills were not limited to the Lower School Library, and we thank our Librarian, Una, for the excellent preparation and confidence she instilled in these young people. All county books were due November 1 and returned by us.

On Tuesday, October 30, we were treated to a symphony by the Youth Philharmonic Orchestra at the New Arlene Schnitzer Performing Arts Center. Students were impressed with the new structure, appreciative of the invitation and enjoyed a morning of diversified music. It was a good trip and we were proud of our kids.

The Famous Person Biography Project is now well underway with most students working on rough drafts at home and maps, portraits, visual projects, etc., at school. Illustrated timelines of each person are finished and soon to be posted. Outlines are completed, and all notes have been checked by us. Students have signed up for a presentation date which is noted on the enclosed signature sheet. In the past, parents have joined us for presentations. This adult participation has been greatly appreciated by former students. With advance notice we try to schedule a specific time to fit your needs, so please join us if you can.

Students and teachers alike are enjoying the transition from fractions to decimals in math classes. By the end of the month, most students will compute decimals with ease and be familiar with converting fractions and decimals. Fractions, an important focus of sixth grade math, are never totally dropped, so do not be concerned if your son or daughter is still hesitant in this area.

Parts of speech have been the focus of language arts classes. Having studied nouns, verbs, and pronouns, we are ready to move on to adjectives, adverbs, and prepositions. Penmanship classes are held on Friday with students having the option of cursive or Italic instruction.

In reading, all groups have finished either Old Yeller or In Search of the Sandhill Crane. Our next two selections are both by John Steinbeck: The Red Pony or The Pearl. To our dismay, there was much confusion with the Biography Book Report which we were never quite able to clear up. We are kicking off the next assignment for a classic book report which will be due December 14. The book report form is enclosed here. Students are expected to have their books selected by November 9th and should spend some time reading nightly.

The permission slip for the trip to Charles Wright Academy, a private school in Tacoma, is enclosed. We will travel by train, departing on November 16th at 8:00 AM. Please return the enclosed permission slip and fare as soon as possible.

Sincerely, \_\_\_\_\_

September 11, 1984

Dear Sixth Grade Parents,

We are off to a good start with sixth grade and looking forward to a full year. We see many bright, questioning minds and a strong desire to put in active days. That's what we love to see--energy.

Back-to-School Night, September 25, is an evening set aside to discuss the year's upcoming activities in the homeroom and special subject classes, and answer questions. We would like to pass along some information, however, which may be helpful through the next few weeks.

1. Homework. We are firm believers that homework reinforces daily skills and helps develop a strong sense of responsibility required by Catlin education. Students failing to complete homework will attend study halls during recesses. Generally, assignments should require 30 minutes to an hour, but this will vary according to individuals. If an assignment demands more time or help than seems reasonable, please encourage your child to return the paper incomplete and let us know. We will pick it up at school. Our intent is to encourage and build confidence with skills, hopefully avoiding evenings of frustration. Via monthly newsletters, you will be informed of major projects and due dates. Occasionally project work will be encouraged at home in lieu of a nightly assignment sheet. Short-term and long-term homework is posted on the blackboard, and students are encouraged to keep an assignment book.
2. Book Reports: Students should always have a silent reading book close at hand. Generally, for a written report, we provide a category, i.e., classic, animal, science fiction, approximately once each month. The first book report, due October 25, is from the biography category.
3. Corrected Papers: Your child's work will be sent home at the end of every month with the exception of creative writing which is kept in a classroom file.
4. Supplies: We have an ample supply of paper and pencils which are set out monthly. If extra supplies are needed, we will let you know.
5. Trips: We are proud to be a "portable" classroom and enjoy spontaneous day jaunts - so be prepared. A written notice will precede trips or activities.

OVER.....

Name: \_\_\_\_\_

Date: Nov. 13, 1984Due Date Nov. 9, 1984Biography on: Kubla Khan

**Oral:** Your introduction was fantastic and grabbed your audience. I have a great deal of respect for your communication skills - especially written. You knew your material well and gave a very smooth delivery. Eye contact with the audience is more important than you give it credit. Work on this for future presentation. You may have been nervous but looked confident.

**Written:** I love your illustrated frontispiece and hope you add the other required information such as grade, school, teachers & date. Also, when is your outline & table of contents. I know joined hand is difficult for you but look on it. Penmanship is neat & legible, however. A few spelling errors could have been picked up with proof reading. Good job on organization of materials.

**Map:** This is a beauty. There is an oriental look which I'm not quite sure how it was achieved. Love your intricate drawing, neat labels.

**Cover:** You are an amazingly detailed artist and our compliments on this lovely & professional looking cover.

**Portrait:** where is the portrait of Mr. Khan?

**Visual:** Did I miss anything? Please mention it.

You obviously learned a great deal about Mr. Khan & this knowledge and interest was radiated. The question and answer period went especially well. In respect the fact that you were able to learn so much about this rather sparsely researched individual. For a first major report, this is pretty good. You must feel proud of your efforts.



Chapter 4  
Classroom Assessment Journals

ABSTRACT

Thirty-two teachers kept journals in which they described the most important of their assessments of students each week for 10 weeks. As a result, nearly 300 classroom assessments were collected for analysis. The results of that analysis are presented in this chapter. First, the frequency of occurrence of various assessment practices is reported. Then major themes gleaned from journal entries are discussed.

## Classroom Assessment Journals

The final component of the Center's current classroom assessment research is based on assessment journals maintained by 12 elementary (K-6) and 20 secondary and middle school (7-12) teachers. The thirty-two journals, each developed over a 10-week period, were intended to aid teachers in tracking and analyzing their own assessment practices. For the assignment, teachers were asked to thoroughly describe the single most important assessment conducted during each week. In doing so, they were to include: a description of the assessment's purpose and importance; the characteristics or areas of knowledge measured; the assessment methods used, such as observation or objective test; the source of the assessment (e.g., teacher-developed or textbook); and the assessment results--how well it worked and what revisions might be needed.

Teachers were urged to include a broad range of assessment options: individual, small group and classroom tests; assessments that measured social characteristics, aptitude and achievement; and measures that included observations, performance tests and objective tests.

### Analysis of Journals

The thirty-two journals include some 290 individual assessment activities. In summarizing these, we focused on two major issues: (1) how teachers described the assessments and their outcomes in respect to purposes, methods and characteristics of the assessment; and (2) what specific issues or perspectives they raised about their own assessment process and their classroom assessment environment. Descriptions are summarized below. We recorded the specific characteristics of elementary and secondary teachers' assessments in order to understand what kinds of assessments were being conducted and what teachers relied on most frequently. Five dimensions of the assessment activity were considered:

- o what purpose teachers most frequently described and how those purposes differed from elementary to secondary programs,
- o what characteristics, such as achievement or aptitude, were being assessed,
- o the assessment strategies used,
- o special dimensions of the test, that is whether it was planned or spontaneous, obtrusive or unobtrusive, and
- o how the assessment was recorded.

### A Composite Assessment Picture

Before describing each dimension separately, we would like to present a composite picture of the assessments described in these teachers' journals. Although the characteristics of teachers' assessments varied considerably across grade level and subject area, teachers' important assessments were strikingly similar.

For elementary teachers, determining students' mastery was the typical assessment purpose. Student achievement rather than social characteristics, aptitude or personality were the key focus of attention, and observation methods were used more frequently than objective tests. In addition, these elementary teachers usually used progress reports rather than grades or anecdotal records to monitor students' progress. This differed from the typical secondary assessment. Secondary teachers were preoccupied with grading. They too focused on student achievement or mastery rather than other personal or social characteristics. Secondary teachers were most likely to use a teacher-developed test, planned in advance, rather than other assessment strategies, and results of the assessment were regularly recorded as a grade in the grade book. The following comments and accompanying tables describe each assessment dimension and teachers' responses more thoroughly.

Purpose. Table 1 summarizes teachers' purposes for conducting assessments. These range from assigning grades to diagnosing, grouping, and evaluating instruction. In their journals, both elementary and secondary teachers focused almost exclusively on three assessment purposes: assigning grades, judging students' mastery of material, and diagnosing individual and group needs. As Table 1 indicates, teachers seldom mentioned other purposes such as sizing up students, grouping, or feedback to parents in these journal entries.

Table 1  
Assessment Purpose\*

|                                     | Elementary |    | Secondary |    | Total |    |
|-------------------------------------|------------|----|-----------|----|-------|----|
|                                     | No.        | %  | No.       | %  | No.   | %  |
| Assign grades                       | 7          | 6  | 73        | 36 | 80    | 25 |
| Diagnose individual and group needs | 26         | 21 | 33        | 16 | 59    | 18 |
| Sizing up                           | 5          | 4  | -         | -  | 5     | 2  |
| Grouping                            | 5          | 4  | 2         | 1  | 7     | 2  |
| Selection for program               | 7          | 5  | 3         | 2  | 10    | 3  |
| Feedback to parents                 | 4          | 3  | 2         | 1  | 6     | 2  |
| Evaluation of instruction           | 3          | 2  | 18        | 9  | 21    | 6  |
| Control and motivate                | 1          | 1  | 7         | 4  | 8     | 2  |
| Feedback to managers                | -          | -  | 2         | 1  | 2     | 1  |
| Communicate expectations            | -          | -  | -         | -  | -     | -  |
| Mastery                             | 67         | 54 | 60        | 30 | 127   | 39 |

In the assessments, teachers most frequent purpose was grading or determining student mastery (64 percent). Diagnosing needs was mentioned in only 18 percent of the journal entries, while only nine percent of secondary entries noted that evaluating instruction was a purpose of the assessment.

\*Some assessments involved multiple purposes.

**Traits Measured.** These teachers were primarily interested in judging student achievement. Eighty-three percent of the assessment entries discussed this characteristic; 17 percent recounted evaluating other characteristics such as social abilities or aptitude. As Table 2 illustrates, the only other characteristic noted with some frequency is social functioning, e.g., interactions with others, etc. Overall, seven percent of the entires, fairly evenly divided between elementary and secondary teachers, discussed assessing this characteristic. Interestingly, very few teachers mentioned assessing critical thinking skills in their journal entries; in fact, few referred to the cognitive level of their test questions despite the emphasis in this course on assessing higher cognitive thinking skills. In judging achievement, most teachers appeared to use a criterion-referenced system; only a few indicated that their assessments compared students to one another (norm-referenced system).

Table 2  
Characteristics

|                                   | Elementary |    | Secondary |    | Total |    |
|-----------------------------------|------------|----|-----------|----|-------|----|
|                                   | No.        | %  | No.       | %  | No.   | %  |
| Achievement mastery (undefined)   | 59         | 48 | 86        | 51 | 145   | 50 |
| Achievement--criterion referenced | 33         | 26 | 46        | 27 | 79    | 27 |
| Achievement--norm referenced      | 9          | 7  | 9         | 5  | 18    | 6  |
| Aptitude                          | 7          | 6  | 4         | 2  | 11    | 4  |
| Higher cognitive functioning      | 1          | 1  | 2         | 1  | 3     | 1  |
| Social characteristics            |            |    |           |    |       |    |
| of individual                     | 11         | 9  | 11        | 6  | 22    | 7  |
| of group                          | -          |    | -         |    | -     |    |
| Personality characteristics       | 1          | 1  | 2         | 2  | 3     | 1  |
| Unspecified                       | 1          | 1  | 2         | 2  | 3     | 1  |
| Other                             | 1          | 1  | 7         | 4  | 8     | 3  |

**Strategies.** In this sample, teacher-developed objective tests, used in 24 percent of the entries, were clearly the single most frequently used assessment method. However, as Table 3 illustrates, behavioral and product observations are the second and third most frequently mentioned areas. When combined together, they occur with greater frequency—in 56 percent of the journal entries—than any other kind of teacher devised assessment. Although a substantial number of elementary teachers also critiqued the use of standardized tests in their journal entries, a far larger number of elementary entries (56 percent) discussed using observation techniques. Few of this group, however, used rating methods or checklists to record their observations.



**Table 3**  
**Assessment Strategies**

|                         | <u>Elementary</u> |          | <u>Secondary</u> |          | <u>Total</u> |          |
|-------------------------|-------------------|----------|------------------|----------|--------------|----------|
|                         | <u>No.</u>        | <u>%</u> | <u>No.</u>       | <u>%</u> | <u>No.</u>   | <u>%</u> |
| Standardized tests      | 18                | 16       | 8                | 4        | 26           | 9        |
| Text embedded           | 12                | 10       | 11               | 6        | 23           | 8        |
| Behavioral observation  | 34                | 30       | 57               | 29       | 91           | 29       |
| Product observation     | 29                | 26       | 52               | 27       | 81           | 26       |
| Teacher developed tests | 20                | 18       | 53               | 28       | 73           | 24       |
| Other                   | -                 |          | 12               | 6        | 12           | 4        |

**Planning of Assessment.** This category attempted to describe whether the assessments were planned in advance similar to most objective tests, or spontaneous, such as an informal observation of a student. In recording information, the categories "planned" and "spontaneous" are used to describe objective tests, while "structured" and "informal" address characteristics of teachers' observations and performance assessments.

The overwhelming majority of objective tests and quizzes described in these journals, were planned rather than spontaneous. In a few instances, however, teachers talked about giving a quick check up quiz to measure students' progress. The structured versus informal continuum also indicated that slightly more teachers described observational assessments that were preplanned (28 percent), such as in judging students' reading skills (this did not necessarily mean that they formally recorded results) rather than used informal assessments. Use of unobtrusive assessment, one of the most useful categories shown in Table 4, illustrates that a small number of teachers conducted an assessment without telling students. Characteristics, such as ability to work cooperatively or to express oneself clearly, were frequently judged via unobtrusive observation by these teachers.

**Table 4**  
**Dimensions**

|                              | <u>Elementary</u> |          | <u>Secondary</u> |          | <u>Total</u> |          |
|------------------------------|-------------------|----------|------------------|----------|--------------|----------|
|                              | <u>No.</u>        | <u>%</u> | <u>No.</u>       | <u>%</u> | <u>No.</u>   | <u>%</u> |
| Objective tests were:        |                   |          |                  |          |              |          |
| Planned                      | 51                | 44       | 69               | 39       | 120          | 41       |
| Spontaneous                  | 3                 | 3        | 9                | 5        | 12           | 4        |
| Performance assessment were: |                   |          |                  |          |              |          |
| Structured                   | 30                | 26       | 52               | 29       | 82           | 28       |
| Informal                     | 19                | 16       | 41               | 23       | 60           | 20       |
| Unobtrusive                  | 11                | 10       | 7                | 4        | 18           | 6        |

**Assessment Record.** Teachers' assessment records, as described in Table 5, indicate that grades in the gradebook were far and away the most frequent type of record maintained by these teachers. A much smaller percentage describe assessments that involved feedback other than a grade. In addition, a substantial number of teachers did not mention their method of recording assessment results, probably because this aspect of the assessment was not specifically requested in the assignment. When teachers described results they usually discussed how well the students did on the test rather than what record they maintained of results. Few teachers maintained any written records or files of students' work apart from a grade in the gradebook.

Table 5  
Form of Record

|                                 | <u>Elementary</u> |          | <u>Secondary</u> |          | <u>Total</u> |          |
|---------------------------------|-------------------|----------|------------------|----------|--------------|----------|
|                                 | <u>No.</u>        | <u>%</u> | <u>No.</u>       | <u>%</u> | <u>No.</u>   | <u>%</u> |
| Grade in book (progress report) | 40                | 38       | 78               | 46       | 118          | 43       |
| Permanent record                | 12                | 11       | 4                | 2        | 16           | 6        |
| Written comment (form)          | 13                | 12       | 19               | 11       | 32           | 11       |
| Verbal comment                  | 8                 | 8        | 17               | 10       | 25           | 9        |
| Non-verbal comment              | 4                 | 4        | 4                | 2        | 8            | 3        |
| Unspecified                     | 23                | 22       | 46               | 27       | 69           | 25       |
| Other                           | 5                 | 5        | 3                | 2        | 8            | 3        |

In summary, the most important weekly assessments described by these elementary and secondary teachers show the following characteristics. The assessment's purpose, regardless of grade level, primarily focused on judging students' achievement or assigning grades (64 percent); only 18 percent of these entries discussed using their assessments to diagnosis student needs. An even smaller percentage indicated that they used the assessment to evaluate instruction. Although teachers' assessment may look at a range of important student characteristics—achievement, social development, aptitude, higher thinking skills, personality, these teachers were unquestionably more concerned with achievement (83 percent) than any other characteristics. This may relate to both the importance of this characteristic for these teachers and to the relative ease in documenting these assessments because they usually involve specific tests.

These teachers showed that they relied on their own observations to conduct a majority of assessments (56 percent). These ranged from evaluating oral reading, writing, map making, to art activities, to group interactions. Teacher-developed objective tests were the second most likely assessments to be described, while text embedded were least likely to be discussed in the journal entries. This, however, may reflect the class' emphasis and the kinds of assessments that are most demanding to conduct rather than these most frequently used. Finally most teachers said they recorded assessment results in the grade book; fewer teachers noted that they provided other kinds of

feedback, either written or verbal, to students. The large number of teachers that did not mention how they record results is as perplexing as the minimal attention teachers gave to discussing or analyzing test results.

This section gave an overview of teachers' journal assessments, noting the purposes, characteristics, strategies, dimensions and records discussed. Teachers' journals also contained rich descriptive information about their experiences in conducting these assessments. The next section describes teachers' perceptions about this aspect of their assessment experiences.

#### Teachers' Analysis and Criticism of Assessments

In describing the assessment activities and environment in their classrooms teachers were asked to comment on the most important assessment they conducted each week in terms of how it worked and how it might be revised in the future. The teachers varied in the degree of reflection and analysis they applied to their weekly assessments. Overall, teachers maintained their own pattern of consistency in writing journal entries, but there was considerable differences in thoroughness across teachers. However, nearly all teachers did comment on these issues.

The following entries demonstrate the variation among teachers in evaluating their assessments. This business teacher ended a full two page typed journal entry with these reactions:

I am most happy to report a success! Perhaps the greatest measure of my results came from the comments made by the students themselves. They were overjoyed with their improved accuracy rates. There was a real showing of pride and accomplishment in a job well done. I was called over by several students to see the improvements they had made.

The results verify what the students already knew. Out of 20 students present for the testing, two students scored an A, eight scored a B, five a C, and five failed.

What a motivator for future assessments! It was a real pleasure to see students pleased with their performance; not to mention the change in attitude I experienced as I corrected their tests. It was no longer a drudgery that I avoided, but a real thrill to see positive results. I certainly plan to continue with similar drills in the future prior to timed writing tests.

In contrast is the physical education teacher, whose entries consisted of one page items about which we gleaned little information, evaluated his assessments in this manner:

This unit was the best of the 3 years. I've taught it before. This type of skill grading works well. I feel the students know where they stand as they play. Sort of a self evaluation. To improve or check this philosophy would be a good idea by having each student write his own grade down based on the posted skills to be mastered.

The total of 290 entries was culled for these evaluative comments on how well assessments worked and how they might be revised in the future. The teachers interpreted the directions freely and frequently wrote incomplete descriptions thereby leaving questions about the classroom assessment environment unanswered. There was, however, enough information in the teachers' responses to draw conclusions with respect to these overarching themes:

- o Are teachers able to analyze and critique assessment processes and outcomes?
- o Do teachers use simple quality control procedures?
- o What is the student's role in assessment?

From the answers to these questions, we were able to identify some general problems with teacher skills in analyzing their own assessments.

To analyze these issues, we grouped our conclusion subissues which the data shed some light upon. Differences between elementary and secondary teachers' perspectives regarding the classroom assessment environment are noted where relevant.

#### ARE TEACHERS ABLE TO ANALYZE AND CRITIQUE THEIR OWN ASSESSMENT PROCESSES AND OUTCOMES?

The overwhelming majority of teachers were able to analyze and critique the assessment process and its outcomes. Teachers were willing and able to point out the weaknesses in their assessment process. In contrast, they were less skilled in providing any indepth analysis of those weaknesses. If an assessment worked, teachers often voiced pride or enthusiasm but they seldom pinpointed factors which contributed to success. When an assessment did not work, teachers usually related it to instructional issues or problems. The following subissues, however, provide a richer understanding of the breadth and scope of their concerns and their readiness to spot shortcomings in the assessment process.

Can teachers critique the instruments used? Teachers were often concerned with the fact that test items used were inadequate in testing the students' knowledge of the unit being assessed. Some were quick to cite the lack of congruity of the instruments to match the purposes of instruction, noting that there was an overemphasis on recall level questions. These observations were often made about district or textbook publisher tests. Their own tests, or tests of their colleagues, were criticized for their excessive wordiness or ambiguity in questions. The need for better planning prior to test construction was often noted.

Can teachers comment on the outcomes of assessments? Frequently teachers evaluated the outcomes in terms of the instructional process. For example, the results often helped them to (a) identify those students needing review



and assistance, (b) conclude that the students needed increased preparation for the test or (c) spot those who needed clearer examples of expectations to improve performance. Teachers also noted the value of increasing the frequency of the evaluation; several teachers noticed the value of immediate feedback.

A few elementary teachers recognized that teaching test taking skills improved students' test results, and they spent considerable time emphasizing this in both their instructional goals and their discussion about their concerns for the assessment. While this is not only an issue in the early grades, it is interesting to note that only the elementary teachers were concerned with testwiseness, particularly as it related to the success of their students on required standardized tests. Additionally, only elementary teachers cited the effects of test anxiety on performance and devised appropriate strategies for dealing with this problem.

Do teachers analyze standardized or district testing practices? Both elementary and secondary teachers spoke forcefully regarding standardized or district testing practices. This was one arena where they were not reticent to point out problems and where they were the most keen in addressing not only the limitations of the practices but the apparent uselessness of standardized test results. Only two teachers mentioned that results from standardized tests were helpful in that they confirmed other data about the children and their own professional perceptions about students' capabilities. Generally, for these teachers, standardized tests failed to provide needed diagnostic information. The tests seemed unrelated to instructional goals in that the content of the tests did not match what they were teaching or what they were expected to teach. Finally, these tests were judged inadequate for evaluating what a "child knows and can do." Teachers gave some of the reasons for this, noting that there are too few items and the tests do not test critical thinking. Elementary teachers recognized that some students have difficulty in following directions on the answer sheets, suggesting the need for more test taking practice. Others found the vocabulary too hard and unrelated to the instructional goals of their classroom. One teacher cynically maintained that 9/10ths of the test data would never be used for instruction and that standardized tests were "given for public relations reasons."

Can teachers use and critique observational or performance assessments? We noted a significant number of teachers who used observational or performance assessments regularly in their classes. Secondary teachers who taught business, physical education, science, art and home economics used performance assessments often and were most likely to have developed more structured rating scales. Sometimes they suggested that they needed to improve the criteria or sharpen the rating scale in order to differentiate student performance more carefully. One home economics teacher observed that simultaneously evaluating both the process and the product was difficult. Where process and product were equally important in cooking she feared inconsistency in her ratings due to the physical limitations of getting around the classroom in time.

The elementary teachers rarely used rating scales or checklists but often noted how much more important observation was than objective testing in assisting the instructional process. They made frequent use of mental notes which were illuminated in the journal but only a few admitted that the reliability of mental notes and snap judgments was highly questionable.

Can teachers analyze problems with grading? We were surprised to find that grading was a continual, nagging problem for secondary teachers but was rarely mentioned as a concern by elementary teachers. This could be attributable to the fact that traditional letter grades are not usually given in the lower primary grades but, rather, teachers prepare progress reports. Secondary teachers struggled constantly with fairness in grading practices, especially as it related to the need for account for social behavior, effort, and attitude. As one art teacher asked rhetorically, "How do I account fairly for effort when there is obviously little talent?" Physical education teachers often wondered how to devise appropriate ways of assessing responsibility, cooperation and effort. Several teachers noted their frustrations with motivating students to complete assignments or take particular assessments seriously when they were ungraded.

Secondary teachers were anxious to communicate how they addressed the grading dilemma, which rarely allowed them to speak to the total learning, strengths/weaknesses or progress of their students. Many of them resorted to some kind of two-level process in assigning grades. One teacher had students grade their own products. A final grade was negotiated when there was a difference in teacher/student perception of performance. Several teachers linked performance ratings with a point system to account for attendance, attitude and effort. An art teacher found a way to resolve the discrepancy between talent and effort by judging student products based upon the level of skill demonstrated in their previous products. One middle school physical education teacher was forceful in his belief that athletic achievement should not be a criterion for grading. He judged students mainly on responsibility, cooperation and effort. A few teachers mentioned the need to check for discrepancies between daily work and letter graded material to arrive at a grade.

#### DO TEACHERS USE SIMPLE QUALITY CONTROL PROCEDURES?

To be clearly understood, this issue must be explored in light of the context within which the journals were written. The teachers who kept the journals were graduate students in a 10-week course in educational measurement. Therefore, through the term, they were receiving new information and guidelines on practical aspects of classroom assessment. This issue focuses on whether or not the training had a discernable impact on teachers' tendency to use simple procedures to maximize assessment reliability and validity.

Do teachers use multiple assessments? Occasionally teachers did recommend the use of more than one assessment technique. This was a significant indication that they wanted to verify initial observations of behavior or

check for the consistency in the results they were obtaining. Teachers either recognized the complex nature of the assessment process or the shortcomings of one particular assessment. They were flexible enough to try another approach.

We were pleased to find a few instances of teachers verifying an initial observation of behavior with another assessment, perhaps a followup quiz or homework check. An elementary teacher did point out that her initial paper and pencil vocabulary test needed a regular observational recheck to see if words were used correctly later. Only one teacher, spurred on by guidelines learned in the measurement class noted that he was planning his first test for higher level thinking skills. In response to her concerns about the limited student data generated from a standardized test, another teacher took extensive anecdotal notes of students' behavior during the testing itself, then used this information with her principal to build a case for the placement of certain individuals.

Some teachers described how they verified judgments of classroom performance with a test activity, or reviewed material and retested to verify mastery. One elementary teacher used a checklist to record specific reading skills, then discussed the results with students to confirm her judgments. A secondary physical education teacher described his pre-post testing process which summarized the data on graphs to communicate improvement quickly to students. While he noted this as time consuming, it was worth the investment as improvement increased and the students stayed motivated.

Do teachers develop clear criteria to judge characteristics specified in a performance assessment? As we noted earlier, these teachers conducted performance assessments regularly. But with the exception of several secondary teachers, few indicated that they use ratings or checklists. Thus, we took special note of those instances where teachers were working to refine or develop their performance criteria checklists and rating scales. Because performance assessment is one of the topics stressed in class, teachers were quick to note their improvement in this area as a direct result of instruction. For instance, after learning how to score essay tests effectively in class, one teacher reported spelling out the assessment criteria more thoroughly and noted a positive impact on student learning. Another teacher clearly specified the criteria for behavioral observations, recognizing that a simple checklist record of behavior may not have been adequate, but was an improvement over the mental record she had been keeping.

Do teachers develop test taking skills in their students? As mentioned earlier, elementary teachers were concerned about their students performing well on standardized tests. A few teachers built in practice time to develop testwiseness. This practice can enhance test validity. One teacher conscientiously prepared her students for the standardized test with practice items. Another evaluated her assessment methods based on the performance of students and their difficulties in mastering the test format. One teacher was so sensitive to the potential negative effects of the testing upon her students performance that she and her colleague provided a snack break and an in-seat art activity break to reduce fatigue during testing.



Do teachers revise assessment to ensure validity? Although teachers often discussed instructional purposes in terms of the entire group, they frequently evaluated their assessment practices in terms of meeting individual needs. A substantial number of teachers made reference to techniques which modified their routine assessment practices to ensure quality. One elementary teacher adjusted the testing mode for non-reading students while another applied a visual technique to vary the form of the test.

Several comments implied that teachers recognized their own ability to be flexible. One teacher noted when an individual assessment may be more appropriate than group assessment, as in the case of a shy student speaking in front of the class. A secondary art teacher pointed out that it is important to plan small goals for students who are completing long term projects. One elementary teacher summarized her assessment goal eloquently when she stated that it was important to identify the conditions when a student functions well and build assessment problems around a student's strengths.

We recorded only a few instances where teachers specified adjustments in the instruction based upon the assessment. Teachers seemed either to take this process for granted or they focused upon meeting individual needs as noted directly above. A secondary physical education teacher stated, "I observe skills as I teach them." Several others used informal and structured performance checks to monitor and adjust instruction. One teacher did say that student frustration meant that reteaching was necessary. Often teachers implied that the verbal and nonverbal cues of students guided instructional decisions such as when to reteach or reassess.

#### WHAT IS THE STUDENTS' ROLE IN ASSESSMENT?

By far the most frequently noted new technique with which teachers experimented was that involving students more actively and directly in the assessment process. Buoyed, perhaps, by the philosophy presented in the measurement course which encouraged teachers to use students as evaluators as much as possible, teachers ventured forth with confidence in this area. They found that increasing student involvement not only made their job easier but was also rich in instructional benefits.

Several teachers noted that student evaluators are highly motivated. The process provided quick feedback to the students. A secondary teacher asked students to note their own level of participation in the unit to help her understand why their performance on the unit test was so poor. One teacher used students' evaluation of the tests to improve instruction because students learned what good performance meant by evaluating it. Another discontinued standardized tests in favor of teacher-made tests based on the reactions and comments of his students when they evaluated results.

Teachers experimented by conducting assessments in teams. One elementary teacher found that using groups to solve problems worked well with students who had partial mastery. In this case, she was using students to provide feedback to teach each other. Generally, teachers believed that engaging students in the assessment process improved the outcome of evaluation by holding the students responsible for their own learning. A few teachers noted, however, that when students were asked to evaluate their own work, they tended to underrate themselves.



One teacher had groups of students developing their own test questions and gave the test to the class. Another elementary teacher fostered collaboration with partners on assignments. Secondary teachers were more likely to have students evaluate their own performance and grade themselves but only one teacher mentioned asking students to explain the process they used in arriving at their ratings.

Cheating was mentioned by several elementary and secondary teachers as a problem which required constant vigilance and forethought. Although no teacher described an actual instance of a violation, many noted it as a reality of classroom life.

#### THE SHORTCOMINGS OF JOURNAL RECORDS

As we read the journals, we noted our own reactions to the teachers' descriptions of their most noteworthy assessment of the week with its failing and promise for improvement in the future. By far our most serious criticism about the entries was that teachers provided little indepth analysis of their assessment and seemed satisfied to simply relay the practice, implying that if the problem were solved the level of analysis was sufficient. When assessments did work, teachers often gauged the "success" of an assessment in terms of student reactions and motivation. For example, teachers said that they "saw student pride and confidence," or, "students did not seem to enjoy this and performance suffered." It appeared that teachers too many of their assessment practices for granted and assumed that others reading the journals or observing their class would intuitively know why something worked without the necessary explanations or justification. When teachers mention that they made judgments based upon observations they usually failed to describe what was being observed by way of the criteria.

Another frequent error was teachers' failure to distinguish the propose of the assessment clearly. For example, teachers confused assessment for mastery with assessment for diagnosis. It appeared that there were often multiple purposes for an assessment but this was overlooked. Related to this was a lack of clarity in relating assessment purposes to the instructional objectives. A few teachers consistently failed to distinguish the assessment from instruction because they critiqued the instructional process, rather than the assessment strategy. For Example, if students did poorly on a test, they were more likely o conclude that more teaching was necessary rather than attributing poor performance to a faulty or unreliable assessment. This may have been due to the fact that many assessments were formative, and teachers were concerned with mastery and the related effectiveness of their teaching.

Criteria for evaluation often was not specified. For example, one art teacher described his performance criteria as "comparing this art project to all artwork in my head." Similarly, few ratings or checklists were used where performance tests were given. Secondary teachers did a more notable job in this area perhaps because they were more likely to present assessments that were summative and for the purposes of grading.

Several teachers consistently disregarded noting how assessment results were used or recorded. This was an easy oversight as the assignment did not specify that this be included. We were left to infer how results were used in many instances based upon the purpose of the assessment. But we learned little about how teachers record this information and whether they distinguish the methods of recording results for themselves or other audiences, such as parents and supervisors.

Summary and Conclusions. For these teachers, mastery of learning was the purpose of assessment, and their assessment methods and concerns related to this purpose. The teachers were eager to critique and analyze the assessment process and its outcomes, but they were less skilled in providing any indepth analysis of particular assessments. They expressed frequent concern for students' growth and well being and were easily attracted to trying strategies which increased student involvement in the assessment process. They were also eager to find new ways to solve the practical problems of classroom assessment. Taken together the assessment techniques teachers described were so varied and situation specific that we marvelled at the size and complexity of the classroom assessment enterprise. It was difficult to determine how much district or collegial support is available to help teachers address assessment concerns but based upon the journal entries along, teachers view standardized, departmental or district testing requirements as less helpful in their daily routine than their own teacher-made assessment techniques. What they need is technical help and support for those in-class assessments. The need help in specifying the criteria for evaluation as well as differentiating the purposes for the assessment. The journal suggest that much can be done to help teachers and improve assessments.

For many of the teachers, however, there was a continual tension between the constraints of time, and the desire to do the best possible job of teaching and assessing students. Some teachers expressed the realization that good instruction and assessment are very complex. However, many teachers also had ideas for new and appropriate strategies for handling complex problems. Yet, others perceived these problems as ongoing and endemic to the nature of the instructional process itself. For instance, when a student was "misplaced" due to parental pressure or faulty use of standardized testing, the teacher had to live with the problem and make the best of the situation. The "solution" was not within her power. Similarly, the presence of learning disabled students taxed the limits of one teacher who brought a special education teacher in to observe her work with these children. No useful solution resulted and the teacher urged that more school or district resources be applied to help teachers gain skills in this area.

## Chapter 5

### Summary of Results

The purpose of these studies of actual classroom assessment environments and journals of key classroom assessment was to document the key ingredients in creating an effective classroom assessment environment. As a result of these studies, we are able to conclude that the assessment environment, like most other dimensions of the classroom, are under the direct control of the teacher. What we have uncovered, however, is a clearer sense of the kinds of factors that constrain the teacher in establishing that environment. There are six sets of constraints.

The most important constraints are those the teacher brings to the classroom. These include knowledge of assessment methodology, prior classroom experience, an array of personal characteristics, perceptions of the students in class, a set of values regarding reasons for and methods of assessment and strategies for communicating expectations and achievement results.

Another set of constraints can arise from characteristics of the classroom itself. These include patterns of staffing and organization as well as the available space and facilities and how they are used.

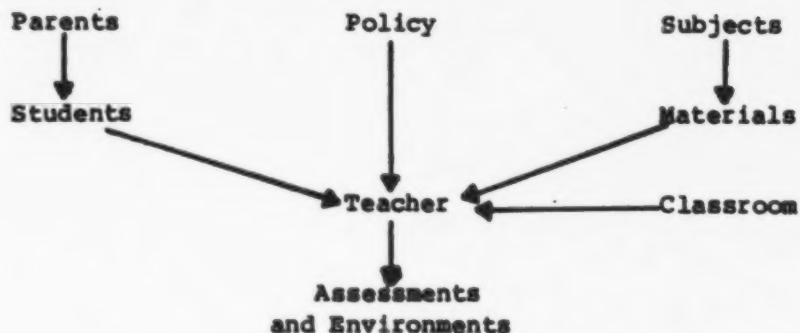
A third set of constraints often come from school and district policy regarding standardized testing, procedures for recording and reporting achievement results, homework requirements, and grouping for special services. Districts manifest specific values with respect to test data that may impact classroom assessment.

Yet another set of constraints can often be found in the texts and materials available to the teacher. Many provide assessment of varying times and varying quality.

Varying characteristics of different school subjects can also influence the nature of the classroom assessment environment. For instance, the perceived importance of the subject can impact the kinds and quality of assessment used. In addition, different kinds of content are more or less amenable to the use of different kinds of evaluation procedures.

Sixth and finally, the assessment environment is defined by the actual assessments used, given these constraining factors. This includes both the actual purposes served and method used to serve those purposes. These assessments are determined by the criteria considered by the teacher in selecting from among the various alternatives available. Those criteria, in turn, contribute greatly to the actual quality of the assessment used.

A graphic representation of the relationships among these factors provides insight into the origins of the assessment environment:



The specific ingredients of these sets of constraining factors—those factors that give rise to the assessment environment in any classroom—are listed and described on the following pages. The descriptions include the underlying continuum along with any factor which can vary for any given classroom and recommended strategies for an observer to use to profile any particular classroom in terms of its assessment environments.



**FACTORS INFLUENCING THE  
CLASSROOM ASSESSMENT ENVIRONMENT**

| <u>Factor</u>                                        | <u>Relevant<br/>Continuum</u>                              | <u>Data<br/>Collection</u>                                                                                |
|------------------------------------------------------|------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|
| 1. Teacher Characteristics                           |                                                            |                                                                                                           |
| A. Knowledge of assessment methodology               |                                                            |                                                                                                           |
| ( 1) Paper and pencil test development               | Well informed-<br>uninformed                               | Test (oral<br>or written)<br>Self report<br>(discussion<br>or interview)<br>Observation of<br>assessments |
| ( 2) Paper and pencil test use                       |                                                            |                                                                                                           |
| ( 3) Performance assessment development              |                                                            |                                                                                                           |
| ( 4) Performance assessment use                      |                                                            |                                                                                                           |
| ( 5) Oral questioning strategies                     |                                                            |                                                                                                           |
| ( 6) Test analysis strategies                        |                                                            |                                                                                                           |
| ( 7) Test score interpretations and use              |                                                            |                                                                                                           |
| ( 8) Grading strategies                              |                                                            |                                                                                                           |
| ( 9) Assessing thinking skills                       |                                                            |                                                                                                           |
| B. Classroom Experience                              |                                                            |                                                                                                           |
| ( 1) In district                                     | Experienced-<br>inexperienced<br>(in years)                | Self report                                                                                               |
| ( 2) In school                                       |                                                            |                                                                                                           |
| ( 3) At grade level                                  |                                                            |                                                                                                           |
| ( 4) With content                                    |                                                            |                                                                                                           |
| C. Personal Characteristics of Teacher               |                                                            |                                                                                                           |
| ( 1) Perceived autonomy<br>in classroom              | Servant of parents/<br>schools-independent<br>professional | Interview                                                                                                 |
| ( 2) Expectations of professional self               | Expect little-<br>expect a lot                             | Intevievw                                                                                                 |
| ( 3) Orientation to classroom structure              | Rigid-flexible                                             | Interview                                                                                                 |
| ( 4) Definition of high quality<br>performance       | Right/wrong-range<br>of quality                            | Observation                                                                                               |
| ( 5) Attention to exceptional student                | Never-frequent                                             | Observatio                                                                                                |
| ( 6) Sense of student norms                          | Clear-unclear                                              | Interview                                                                                                 |
| ( 7) Willingness to experiment with<br>class         | No risks-risk taker                                        | Interview                                                                                                 |
| ( 8) Willingness to experiment with<br>student       | No risks-risk taker                                        | Interview                                                                                                 |
| ( 9) Orientation to class                            | One on one-group                                           | Observation                                                                                               |
| (10) Expectations of working<br>relationships        | Cooperation-<br>competition                                | Interview,<br>observation                                                                                 |
| (11) Attributions of success/<br>failure of students | Student responsible-<br>teacher responsible                | Interview                                                                                                 |
| (12) Orientation to punctuality                      | Demand it-unconcerned                                      | Observation                                                                                               |
| (13) Definition of "on task"                         |                                                            | Interview                                                                                                 |

| <u>Factor</u>                              | <u>Relevant<br/>Continuum</u>                                               | <u>Data<br/>Collection</u>                         |
|--------------------------------------------|-----------------------------------------------------------------------------|----------------------------------------------------|
| D. Teacher's Perceptions of Current Class  |                                                                             |                                                    |
| ( 1) Ability to learn                      | Low-high                                                                    | Interview,<br>discussion                           |
| ( 2) Variation in ability                  |                                                                             |                                                    |
| ( 3) Rate of achievement                   | Low-high,<br>Accelerating-<br>deceleration                                  |                                                    |
| ( 4) Variation in rate                     | High-Low                                                                    |                                                    |
| ( 5) Willingness to learn                  |                                                                             |                                                    |
| ( 6) Variation in willingness              |                                                                             |                                                    |
| ( 7) Maturity                              | Responsible-<br>irresponsible                                               |                                                    |
| ( 8) Study skills                          | Developed-undeveloped<br>Follow directions-<br>have difficulty with<br>same |                                                    |
| ( 9) Social skills                         | Developed-undeveloped<br>Cooperative-disruptive<br>Complaint-resistant      |                                                    |
| (10) Willingness to perform                | Willing-reticent<br>Passive-aggressive                                      |                                                    |
| (11) Gender differences                    | Important-unimportant<br>Related interactions?                              | Observation                                        |
| (12) Feedback needs                        | Frequent-infrequent<br>Individual-group<br>Verbal-models                    |                                                    |
| (13) Self-assessment skills                | Developed-undeveloped                                                       |                                                    |
| (14) Student sense of what's fair          | Clear-unclear                                                               |                                                    |
| (15) Reactions to testing                  | Positive-negative<br>Anxious-tranquil                                       |                                                    |
| (16) Parental expectations                 | Clear-unclear<br>High-low                                                   |                                                    |
| E. Valued reasons for assessment           |                                                                             |                                                    |
| ( 1) Diagnosing group needs                | Important-<br>unimportant                                                   | Interview with<br>anecdotes<br>from<br>observation |
| ( 2) Diagnosing individual needs           |                                                                             |                                                    |
| ( 3) Sizing students up in fall            |                                                                             |                                                    |
| ( 4) Selecting for special services        |                                                                             |                                                    |
| ( 5) Controlling students                  |                                                                             |                                                    |
| ( 6) Motivating students                   |                                                                             |                                                    |
| ( 7) Evaluating instruction                |                                                                             |                                                    |
| ( 8) Communicating academic expectations   |                                                                             |                                                    |
| ( 9) Communicating behavioral expectations |                                                                             |                                                    |
| (10) As test taking training for students  |                                                                             |                                                    |

| <u>Factor</u>                                              | <u>Relevant<br/>Continuum</u> | <u>Data<br/>Collection</u> |
|------------------------------------------------------------|-------------------------------|----------------------------|
| <b>F. Valued Assessment Methods</b>                        |                               |                            |
| ( 1) Daily written assignments                             | Useful-useless                | Interview                  |
| ( 2) Observation and judgment                              |                               |                            |
| ( 3) Paper and pencil tests                                |                               |                            |
| ( 4) Assessments from text                                 |                               |                            |
| ( 5) Assessments from other teachers                       |                               |                            |
| ( 6) Oral recitation in class                              |                               |                            |
| ( 7) Standardized tests                                    |                               |                            |
| ( 8) Student peer assessment                               |                               |                            |
| ( 9) Student self assessment                               |                               |                            |
| (10) Group assessments                                     |                               |                            |
| <b>G. Valued Strategies for Communicating Expectations</b> |                               |                            |
| ( 1) Written, verbal                                       | Useful-useless                | Interview                  |
| ( 2) Oral                                                  |                               |                            |
| ( 3) Via model or example                                  |                               |                            |
| ( 4) Via assessments                                       |                               |                            |
| <b>H. Strategies for Providing Feedback to</b>             |                               |                            |
| ( 1) Students                                              |                               | Observation                |
| (a) formality                                              | Formal-informal               |                            |
| (b) mode                                                   | Written-oral                  |                            |
| (c) frequency                                              | Never-continuously            |                            |
| (d) form                                                   | Grades-comments               |                            |
|                                                            | Private-public                |                            |
| (e) focus                                                  | Achievement?                  |                            |
|                                                            | Ability?                      |                            |
|                                                            | Social personal traits?       |                            |
| ( 2) Parents                                               |                               |                            |
| (a) formality                                              | Formal-informal               |                            |
| (b) mode                                                   | Written-oral                  |                            |
| (c) frequency                                              | Never-continuously            |                            |
| (d) form                                                   | Grades-comments               |                            |
|                                                            | Private-public                |                            |
| (e) focus                                                  | Achievement?                  |                            |
|                                                            | Ability?                      |                            |
|                                                            | Social personal traits?       |                            |
| ( 3) Supervisors                                           |                               |                            |
| (a) formality                                              | Formal-informal               |                            |
| (b) mode                                                   | Written-oral                  |                            |
| (c) frequency                                              | Never-continuously            |                            |
| (d) form                                                   | Grades-comments               |                            |
|                                                            | Private-public                |                            |
| (e) focus                                                  | Achievement?                  |                            |
|                                                            | Ability?                      |                            |
|                                                            | Social personal traits?       |                            |

| <u>Factor</u>                                           | <u>Relevant<br/>Continuum</u> | <u>Data<br/>Collection</u>                      |
|---------------------------------------------------------|-------------------------------|-------------------------------------------------|
| <b>2. Classroom Characteristics</b>                     |                               |                                                 |
| <b>A. Staffing and Organization</b>                     |                               | Observation                                     |
| ( 1) Teachers present                                   | Alone-team                    |                                                 |
| ( 2) Teacher aid                                        | Available?                    |                                                 |
| ( 3) Parental assistance                                | Never-frequently              |                                                 |
| ( 4) Free planning periods                              | None-many                     |                                                 |
| <b>B. Use of Space</b>                                  |                               |                                                 |
| ( 1) Density of students (space/student)                | Much-little                   | Observation                                     |
| ( 2) Physical arrangement                               | Describe it                   |                                                 |
| ( 3) Assessment displays                                |                               |                                                 |
| (a) records of achievement                              | Present?                      |                                                 |
| (b) models of good work                                 |                               |                                                 |
| ( 4) Neatness and order criteria                        | Explicit-implicit             |                                                 |
| ( 5) Strategies for noise control                       | Present-absent                |                                                 |
| during assessment                                       | Rigorously applied?           |                                                 |
| (a) specific rules                                      |                               |                                                 |
| (b) seating arrangements                                |                               |                                                 |
| <b>C. Support for Assessment</b>                        |                               | Observation                                     |
| ( 1) Resources such as library,<br>movies, etc.         | Available?<br>Used?           |                                                 |
| ( 2) Equipment such as size lab                         | Available?<br>Used?           |                                                 |
| ( 3) Facilities such as computers<br>and overflow space | Available?<br>Used?           |                                                 |
| <b>3. School and District Policy</b>                    |                               | Overt policy-<br>see policy<br>manuals          |
| <b>A. Standardized Testing Policy</b>                   |                               | Covert policy-<br>interview ad-<br>ministrators |
| ( 1) Sense of accountability for scores                 | Strong-weak                   |                                                 |
| ( 2) Scores valued and used                             | Useful-useless                |                                                 |
| ( 3) Time committed to testing                          | (In hours)                    |                                                 |
| ( 4) Perceive importance of<br>testing experience       | Important-unimportant         |                                                 |
| <b>B. Policy Regarding Record Keeping and Reporting</b> |                               |                                                 |
| ( 1) Frequency                                          | Frequent-infrequent           |                                                 |



| <u>Factor</u>                                                                        | <u>Relevant<br/>Continuum</u>                                     | <u>Data<br/>Collection</u>                 |
|--------------------------------------------------------------------------------------|-------------------------------------------------------------------|--------------------------------------------|
| ( 2) Content                                                                         | Achievement?<br>Ability?                                          |                                            |
| ( 3) Form                                                                            | Social personal traits?<br>Standard-individual<br>Grades-comments |                                            |
| ( 4) Target of reports                                                               | Students?<br>Parents?<br>Supervisors?                             |                                            |
| C. Policy Regarding Homework                                                         |                                                                   |                                            |
| ( 1) Frequency                                                                       | Intervals or<br>frequency specified?                              |                                            |
| ( 2) Form of homework required                                                       | Written only?                                                     |                                            |
| ( 3) Grading policy                                                                  | % cutoffs?                                                        |                                            |
| ( 4) Time appropriation                                                              | Specified?                                                        |                                            |
| ( 5) Reliance on text assignments                                                    | Required?                                                         |                                            |
| D. Policy Regarding Grouping for Special Services<br>(L.D., Gifted, Chapter 1, etc.) |                                                                   |                                            |
| ( 1) Data requirements                                                               | Clear-unclear                                                     |                                            |
| ( 2) Classroom data allowed                                                          | Grades-comments                                                   |                                            |
| ( 3) Criteria for selection                                                          | Clear-unclear                                                     |                                            |
| E. Valued District Uses of Test Data                                                 |                                                                   |                                            |
| ( 1) To establish teacher accountability                                             | Important-unimportant                                             |                                            |
| ( 2) To compare schools, classes, etc.                                               |                                                                   |                                            |
| ( 3) To show achievement trends                                                      |                                                                   |                                            |
| 4. Characteristics of Texts and Materials                                            |                                                                   |                                            |
| A. Assessments Provided in Texts<br>(may vary with subject)                          |                                                                   | Examine tests<br>& associated<br>materials |
| ( 1) Discussion questions for class<br>recitation                                    | Present-absent                                                    |                                            |
| ( 2) Homework assignments in text                                                    |                                                                   |                                            |
| ( 3) Workbook study sheets                                                           |                                                                   |                                            |
| ( 4) Assessment guidelines in<br>teacher's guide                                     |                                                                   |                                            |
| ( 5) Paper and pencil tests                                                          |                                                                   |                                            |
| ( 6) Performance assessments                                                         |                                                                   |                                            |

| <u>Factor</u>                                                           | <u>Relevant<br/>Continuum</u>                                        | <u>Data<br/>Collection</u> |
|-------------------------------------------------------------------------|----------------------------------------------------------------------|----------------------------|
| B. Quality of Assessments Offered                                       |                                                                      |                            |
| ( 1) Validity                                                           |                                                                      |                            |
| (a) match content of text                                               | Miss-match                                                           |                            |
| (b) match cognitive levels of<br>text and recitation                    |                                                                      |                            |
| (c) sample representatively?                                            |                                                                      |                            |
| ( 2) Reliability                                                        |                                                                      |                            |
| (a) length                                                              | Appropriate-inappropriate                                            |                            |
| (b) methods                                                             | Objective-subjective                                                 |                            |
| (c) item construction                                                   | Clear-unclear                                                        |                            |
| (d) scoring guidelines                                                  | Absent-detailed                                                      |                            |
| ( 3) Ease of use                                                        | Convenient-inconvenient                                              |                            |
| C. Nontext Materials Used in Assessment? (describe them)                |                                                                      |                            |
| 5. Characteristics of School Subject                                    |                                                                      |                            |
| A. Perceived Importance of Subject as Seen by<br>(will vary by subject) |                                                                      |                            |
| ( 1) Students                                                           | Important-unimportant                                                | Interview                  |
| ( 2) Parents                                                            |                                                                      |                            |
| ( 3) Teachers                                                           |                                                                      |                            |
| ( 4) School                                                             |                                                                      |                            |
| ( 5) District                                                           |                                                                      |                            |
| B. Other Indicators of Importance<br>(by subject)                       | Time allotted<br>Required-optional                                   | Examine<br>records         |
| C. Relationship of Content to Assessment<br>Options (by subject)        |                                                                      | Examine<br>content         |
| ( 1) Written assignments                                                | Amenable to content-                                                 |                            |
| ( 2) Teacher observation and judgment                                   | not amenable                                                         |                            |
| ( 3) Classroom paper and pencil tests                                   |                                                                      |                            |
| ( 4) Assessments from texts                                             |                                                                      |                            |
| ( 5) Assessments from other teachers                                    |                                                                      |                            |
| ( 6) Oral recitation                                                    |                                                                      |                            |
| ( 7) Self assessments                                                   |                                                                      |                            |
| ( 8) Peer assessments                                                   |                                                                      |                            |
| ( 9) Standardized tests                                                 |                                                                      |                            |
| (10) Group assessment                                                   |                                                                      |                            |
| (11) Application of rules of<br>evidence                                | Punctual completion-<br>quality<br>Right/wrong-degrees<br>of quality |                            |

| <u>Factor</u>                                                         | <u>Relevant<br/>Continuum</u> | <u>Data<br/>Collection</u> |
|-----------------------------------------------------------------------|-------------------------------|----------------------------|
| <b>6. Characteristics of Assessments</b>                              |                               |                            |
| <b>A. Assessment Purposes Demonstrated</b><br>(may vary with subject) |                               | Observation                |
| ( 1) Diagnosing group needs                                           | Never used-                   |                            |
| ( 2) Diagnosing individual needs                                      | frequently used               |                            |
| ( 3) Sizing students up in fall                                       |                               |                            |
| ( 4) Selecting for special services                                   |                               |                            |
| ( 5) Controlling students                                             |                               |                            |
| ( 6) Motivating students                                              |                               |                            |
| ( 7) Evaluating instruction                                           |                               |                            |
| ( 8) Communicating academic expectations                              |                               |                            |
| ( 9) Communicating behavioral expectations                            |                               |                            |
| (10) As test taking training for students                             |                               |                            |
| <b>B. Assessment Practices Used</b><br>(may vary with subject)        |                               | Observation                |
| ( 1) Daily written assignments                                        | Never used-                   |                            |
| ( 2) Observation and judgment                                         | frequently used               |                            |
| ( 3) Paper and pencil tests                                           |                               |                            |
| ( 4) Assessments from text                                            |                               |                            |
| ( 5) Assessments from other teachers                                  |                               |                            |
| ( 6) Oral recitation in class                                         |                               |                            |
| ( 7) Standardized tests                                               |                               |                            |
| ( 8) Student peer assessment                                          |                               |                            |
| ( 9) Student self assessment                                          |                               |                            |
| (10) Group assessments                                                |                               |                            |
| <b>C. Categories of Thinking Skills Addressed</b>                     |                               | Observation                |
| ( 1) Levels defined                                                   | No-clearly                    |                            |
| ( 2) Levels in assessment match<br>instruction                        | Miss-match                    |                            |
| ( 3) Describe levels used                                             |                               |                            |
| <b>D. Criteria Used by Teacher in Selecting Method</b>                |                               | Interview                  |
| ( 1) Fit of results to purpose                                        | Important-unimportant         |                            |
| ( 2) Match to material taught                                         | Evaluated how?                |                            |
| ( 3) Ease of development (easy access)                                |                               |                            |
| ( 4) Ease of scoring                                                  |                               |                            |
| ( 5) Origin of test                                                   |                               |                            |
| ( 6) Time required to administer                                      |                               |                            |
| ( 7) Degree of objectivity                                            |                               |                            |
| ( 8) Test security                                                    |                               |                            |

| <u>Factor</u>                 | <u>Relevant<br/>Continuum</u> | <u>Data<br/>Collection</u> |
|-------------------------------|-------------------------------|----------------------------|
| E. Quality of Assessments     |                               | Examine<br>assessments     |
| ( 1) Validity                 |                               |                            |
| (a) match to content          | Miss-match                    |                            |
| (b) match to cognitive levels |                               |                            |
| (c) representative sample?    |                               |                            |
| ( 2) Reliability              |                               |                            |
| (a) length                    | Appropriate-inappropriate     |                            |
| (b) methods                   | Objective-subjective          |                            |
| (c) item construction         | Clear-unclear                 |                            |
| (d) scoring procedures        | Appropriate-inappropriate     |                            |

Note: The implication of these many factors for further research and teacher training in assessment will be explored in the final report.